MULTIVARIATE LINEAR PREDICTIVE SPECTRAL ANALYSIS EMPLOYING WEIGHTED FORWARD AND BACKWARD AVERAGING

A GENERALIZATION OF BURG'S ALGORITHM

NAVAL UNDERWATER SYSTEMS CENTER,
NEW LONDON, CONNECTICUT

13 OCTOBER 1976

NUSC Technical Report 5501

NUSC Technical Report 5501

Multivariate Linear Predictive Spectral Analysis Employing Weighted Forward and Backward Averaging: A Generalization of Burg's Algorithm

Albert H. Nuttall
Advanced Systems Technology

13 October 1976

NUSC
NAVAL UNDERWATER SYSTEMS CENTER
Newport, Rhode Island • New London, Connecticut

PREFACE

This research was conducted under NUSC Project No. A-752-05, "Applications of Statistical Communication Theory to Acoustic Signal Processing," Principal Investigator, Dr. A. H. Nuttall (Code 313), and Navy Project No ZR000 01, Program Manager, T. A. Kleback (MAT 03521), and under NUSC Project No. A-600-00, "Acoustic Communications for Submarines and Surface Vessels," Principal Investigator, A. W. Ellinthorpe (Code 3103), Program Managers, J. Calabrese (SEA 660E-31) and H. Bezdek (ONR 486).

The author would like to acknowledge several helpful discussions held with Dr. Philip Abraham (the Technical Reviewer for this report) and with Dr. Edward Eby during the progress of this investigation.

REVIEWED AND APPROVED: 13 October 1976

W. A. Von Winkle
Head: Special Projects Department

The author of this report is located at the New London Laboratory, Naval Underwater Systems Center, New London, Connecticut 06320.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TR 5501	2. GOVT ACCESSION NO	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MULTIVARIATE LINEAR PREDICTIVE SPECTRAL ANALYSIS EMPLOYING WEIGHTED FORWARD AND BACKWARD AVERAGING: A GENERALIZATION OF BURG'S ALGORITHM		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Albert H. Nuttall		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Underwater Systems Center New London Laboratory New London, CT 06320		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS A-752-05/A-600-00 ZR000 01
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Sea Systems Command Washington, DC 20362		12. REPORT DATE 13 Oct 1976 13. NUMBER OF PAGES ~~90~~ 100
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Spectral Analysis
Multivariate Process
Linear Predictive Filter
Forward and Backward Averaging
Generalization of Burg's Algorithm
Maximum Entropy
Weighted Error Matrices

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A method for multivariate linear predictive spectral analysis, employing weighted forward and backward averaging, is presented and programmed in FORTRAN. The method constitutes a generalization of Burg's univariate algorithm to the multivariate case. The essential analytical procedure is to minimize the trace of the sum of the weighted forward and backward error matrices by choice of the partial correlation coefficients, subject to a linear matrix constraint which guarantees that the forward-extrapolated and

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

backward-extrapolated correlation matrix estimates are Hermitians of each other. The choice of error weighting is important and is discussed. Solution of a bilinear matrix equation is required in the algorithm.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

LIST OF SYMBOLS

t	Time
X, X(t), X_n	Multivariate stationary, zero-mean, random process eq. (1)
M	Dimensionality of X, X(t), X_n
Δ	Common sampling interval in time
R_k	Correlation matrix of input process $\{X_n\}$ at delay $k\Delta$, eq. (2)
Overbar	Ensemble average
Superscript T	Transpose
Superscript *	Conjugate
Superscript H	Conjugate Transpose
N	Number of M-dimensional data samples available
f	Frequency (Hz)
i	$\sqrt{-1}$
G(f)	Spectrum of $\{X_n\}$, eq. (3)
p	Range of values of k for which R_k is known; order of prediction filter
$\hat{X}_k$	Forward predicted value of X_k, eq. (4)

LIST OF SYMBOLS

$A_n, A_n^{(p)}$	n-th coefficient of forward predictive filter, eq. (4)
Y_k, Z_k	Instantaneous error, eqs. (5) and (9)
tr	Trace of a square matrix
$\check{X}_{k-p}$	Backward predicted value of X_{k-p}, eq. (8)
$B_n, B_n^{(p)}$	n-th coefficient of backward predictive filter, eq. (8)
U_p, V_p	Correlation matrices of residuals, eqs. (12) and (30)
C_{p-1}, D_{p-1}	Auxiliary matrices, eqs. (13), (35), and (36)
$Y_k^{(p)}, Z_k^{(p)}$	p-th order forward and backward residuals, eq. (17)
I	Identity matrix
$A_p^{(p)}, B_p^{(p)}$	Partial correlation coefficients
z	Variable of z-transforms
$\mathcal{H}_A^{(p)}(z), \mathcal{H}_B^{(p)}(z)$	Transfer functions (z-transforms) of p-th stage, eq. (23)
D	Arbitrary MxM matrix
$\tilde{X}_n$	Scaled process, eq. (24)
$\tilde{R}_k$	Correlation of scaled process, eq. (25)
$\tilde{A}_n^{(p)}$	n-th coefficient for scaled process, eq. (28)
det	Determinant of a square matrix
q, q_n	Arbitrary Mx1 complex matrix
$F_m^{(p)}$	Crosscorrelation of residual and input, eq. (38)
$\hat{R}_m^{(p)}$	Forward extrapolated correlation matrix, eq. (52)
$\check{R}_m^{(p)}$	Backward extrapolated correlation matrix, eq. (55)
$G^{(p)}(f)$	p-th order spectral approximation, eqs. (67) and (69)
$H_A^{(p)}(f), H_B^{(p)}(f)$	T ansfer functions of p-th stage, eqs. (68) and (70)
W_k	White noise, eq. (73)

LIST OF SYMBOLS

δ_{m0}	Kronecker delta: 1 if m = 0; 0 otherwise
$K_m^{(p-1)}$	Forward extrapolated correlation estimate, eq. (86)
$R_{-m}^{(p-1)}$	Backward extrapolated correlation estimate, eq (92)
V_{p-1}	Estimate of the correlation of backward residual at zero time delay, eq. (89)
U_{p-1}	Estimate of the correlation of forward residual at zero time delay, eq. (95)
S_{mn}	Correlation matrix, eq. (103)
$Y_k^{(p)}, Z_k^{(p)}$	Errors (residuals), eq. (111)
E_p, F_p	Error matrices, eqs. (112) and (115)
$S_{p-1}^{(\ )}$	Auxiliary correlation matrices, eq. (114)
Λ_a	Weighting matrix, eq. (117)
Λ	Hermitian non-negative matrix = $\Lambda_a^H \Lambda_a$
$\Lambda_{p-1}, \Gamma_{p-1}$	Weighting matrices, eq. (121)
G_p	Auxiliary matrix, eq. (122)
α, β, μ, ν	Auxiliary matrices, eqs. (126) and (127)
ψ_{p-1}, Ψ_{p-1}	Square root matrices, eq. (142)
$\tilde{Y}_k^{(p)}, \tilde{Z}_k^{(p-1)}$	Scaled processes, eq. (143)
$\tilde{\mathcal{Y}}_k^{(p)}, \tilde{\mathcal{Z}}_k^{(p)}$	Error processes, eqs. (144) and (148)
$\tilde{A}_p^{(p)}, \tilde{B}_p^{(p)}$	Scaled partial correlation coefficients, eqs. (145) and (149)
$\mathcal{E}_p, \mathcal{F}_p$	Estimated correlations of errors, eqs. (146) and (150)
a_k, A_k	Auxiliary quantities, eqs. (160) and (161)
$G^{(p)}(f)$	Spectral estimate, eqs. (165) and (167)
Adj	Adjoint of a square matrix
FFT	Fast Fourier transform

LIST OF SYMBOLS

N_F	Number of frequency cells in range $\left(-\frac{1}{2\Delta}, \frac{1}{2\Delta}\right)$
$R_A(f), I_A(f)$	Real and imaginary parts of Adjoint, eq. (172)
$M(f)$	Auxiliary matrix, eqs. (174) and (179)
$XX(f), YY(f)$	Real and imaginary parts of $H_A^{(p)}(f)$, eq. (175)
$XX_A(f), YY_A(f)$	Adjoints, eqs. (176) and (177)
AIC_p	Akaike Information Criterion, eq. (180)
$\ln$	Natural logarithm
p_{max}	Maximum value of p considered
p_{best}	Best order of p for spectral estimate

INTRODUCTION

Spectral analysis of stationary random processes via linear predictive, maximum entropy, and autoregressive techniques has attracted much attention lately, especially for short data segments; see, for example, the bibliographies listed in references 1, 2, and 3. For a univariate process, it appears that the Burg algorithm (Ref. 4), which guarantees a stable correlation recursion, is as good as any of the currently available techniques of similar nature that employ an all-pole model of the available process (Ref. 3).

Accordingly, it is desirable to develop a spectral analysis technique for the multivariate case in such a way that: we employ a physically meaningful error minimization for the determination of the filter coefficients; the technique yields a stable correlation recursion; and it reduces to Burg's algorithm for the univariate case. It will be shown in the following that we have accomplished these goals, with the exception that we have not proved (or disproved) the stability requirement. A FORTRAN program for this spectral analysis technique was published in Ref. 5, along with an example of its application. Virtually simultaneously, the same technique was investigated independently and published in Ref. 6. In this report, we will document the derivations and equations that lead to the program presented in Ref. 5, and indicate an extension of that result.

Our approach in this report will be to investigate, in some detail, first the case where the correlation of the multivariate process under consideration is known for a limited range of argument values, and to extract all the relevant important properties of the solution so that they may be forced to be satisfied later when we treat the unknown correlation case. This property-

extraction procedure will be found to: furnish guides to the analysis of the unknown correlation case; allow us to cut down on computer execution time and storage by employing the properties; and make us aware of some of the shortcomings of the unknown (versus known) correlation cases. This procedure should also be helpful to those who are not thoroughly familiar with spectral analysis of multivariate processes and their properties.

Throughout this report, we assume we are dealing with equispaced samples of a stationary zero-mean complex random process X(t) of dimensionality M; that is, sample

$$X(n\Delta) \equiv X_n \equiv \left[x_n^{(1)} \cdots x_n^{(M)}\right]^T \tag{1}$$

is an M x 1 column matrix, where Δ is the common sampling interval for all the component processes of X(t). It is not assumed that X(t) is Gaussian.

In section 2, we will assume that the correlation matrix of process $\{X_n\}$, namely the M x M matrix*

$$R_k \equiv \overline{X_n X_{n-k}^H} = R_{-k}^H \tag{2}$$

is known exactly for a limited range of values of k, and will show how an approximation for the spectrum of process $\{X_n\}$ can be obtained. In section 3, the input correlation matrix R_k will be unknown, and all that is available is a finite set of N data samples, $X_1, X_2, \ldots, X_N$, from which an estimate of the spectrum of process $\{X_n\}$ is desired. The end result will be a FORTRAN program for multivariate spectral analysis.

*The case of complex samples is treated so that we can handle complex envelope or complex demodulated processes. Specialization to real processes is immediate, and (2) becomes $R_k = R_{-k}^T$. An overbar indicates an ensemble average, superscript T denotes a transpose, and superscript H denotes a conjugate transpose. Matrices are indicated by capital letters.

2. KNOWN CORRELATION

If the correlation in (2) is known for all k, the standard (double-sided) definition of the spectrum of process $\{X_n\}$ is

$$G(f) = \Delta \sum_{k=-\infty}^{\infty} \exp(-i2\pi f k\Delta)\, R_k \ , \quad |f| < \frac{1}{2\Delta} \ . \tag{3}$$

The complex M x M matrix G(f) is Hermitian and non-negative definite for any value of frequency f (see appendix A), but need not be even in frequency f. When R_k is not known for all k, but only for a range $|k| \le p$, an approximation to (3) must be accepted; this problem will be pursued below.

2.1 DERIVATION OF EQUATIONS

Suppose M-dimensional samples $X_{k-p}, \ldots, X_{k-1}$ are available, and we attempt a one-step linear prediction of X_k according to the p-th order operation

$$\hat{X}_k \equiv \sum_{n=1}^{p} A_n X_{k-n} \ , \tag{4}$$

where complex coefficient matrix A_n is M x M, n = 1, 2, ..., p. The instantaneous error at time $k\Delta$ is defined as

$$Y_k = X_k - \hat{X}_k = -\sum_{n=0}^{p} A_n X_{k-n} \ , \quad A_0 \equiv -I. \tag{5}$$

The linear operators in (4) and (5) constitute stable linear filters regardless of the choice of coefficients; the filter of (4) is called the predictive filter, that of (5) is called the predictive error filter. Notice that we are not assuming that process $\{X_n\}$ actually satisfies an autoregressive relation; rather we are simply attempting to linearly predict $\{X_n\}$ on the basis of the most recent p past values.

The minimum value of the scalar error

$$\overline{\underline{Y}_k^H \underline{Y}_k} = \mathrm{tr}\, \overline{Y_k Y_k^H} \quad , \tag{6}$$

by choice of coefficients $\{A_n\}_1^p$, is given (in appendix B) by the solution of the linear matrix equations

$$\sum_{n=1}^{p} A_n^{(p)} R_{m-n} = R_m \quad , \quad 1 \le m \le p, \tag{7}$$

where the explicit dependence on the order p is indicated. Knowledge of R_k for $|k| \le p$ is required in (7).

Before we discuss the solution of (7) for $\{A_n^{(p)}\}_1^p$, we consider one-step linear "backward prediction" of process $\{X_n\}$. Suppose samples X_k, X_{k-1}, ..., X_{k-p+1} are available, and we attempt a one-step linear prediction of X_{k-p} according to

$$\check{X}_{k-p} \equiv \sum_{n=1}^{p} B_n X_{k-p+n} \quad . \tag{8}$$

The instantaneous error is defined as

$$Z_k \equiv X_{k-p} - \check{X}_{k-p} = - \sum_{n=0}^{p} B_n X_{k-p+n} \quad , \quad B_0 \equiv -I. \tag{9}$$

The minimum value of the scalar error

$$\overline{\underline{Z}_k^H \underline{Z}_k} = \mathrm{tr}\, \overline{Z_k Z_k^H} \quad , \tag{10}$$

by choice of coefficients $\{B_n\}_1^p$, may be shown (in a manner similar to that of appendix B) to be given by the solution of the linear matrix equations

$$\sum_{n=1}^{p} B_n^{(p)} R_{n-m} = R_{-m} \quad , \quad 1 \le m \le p. \tag{11}$$

For the optimum coefficients in (7) and (11), we find (see appendix B) that the optimum error matrices take the form

$$\text{opt } \overline{Y_k Y_k^H} = R_0 - \sum_{n=1}^{p} A_n^{(p)} R_n \equiv U_p , \quad U_0 \equiv R_0 ,$$
$$\text{opt } \overline{Z_k Z_k^H} = R_0 - \sum_{n=1}^{p} B_n^{(p)} R_n \equiv V_p , \quad V_0 \equiv R_0 . \tag{12}$$

In general, these two matrices, their diagonal elements, and their traces are unequal (as the simple example of p=1 will show). However, their determinants are equal, as will be shown in subsection 2.2.

The solutions of (7) and (11) can be accomplished simultaneously in a recursive fashion (Ref. 7). Define

$$C_{p-1} \equiv - \sum_{n=0}^{p-1} A_n^{(p-1)} R_{p-n} , \qquad A_0^{(p)} \equiv -I ,$$
$$D_{p-1} \equiv - \sum_{n=0}^{p-1} B_n^{(p-1)} R_{n-p} , \qquad B_0^{(p)} \equiv -I . \tag{13}$$

Then

$$A_p^{(p)} = C_{p-1} V_{p-1}^{-1} , \quad B_p^{(p)} = D_{p-1} U_{p-1}^{-1} \tag{14}$$

and

$$\left.\begin{aligned} A_n^{(p)} &= A_n^{(p-1)} - A_p^{(p)} B_{p-n}^{(p-1)} \\ B_n^{(p)} &= B_n^{(p-1)} - B_p^{(p)} A_{p-n}^{(p-1)} \end{aligned}\right\} \quad 1 \le n \le p-1 \qquad (p \ge 2). \tag{15}$$

These relations will be simplified somewhat in subsection 2.2. For M=1, a univariate process, (7) and (11) immediately yield

$$A_n^{(p)} = B_n^{(p)*} \quad \text{for} \quad M = 1, \tag{16}$$

where we have used (2) in the form $R_k = R_{-k}^*$ for a univariate process. No such simple relation as (16) holds for $M \geq 2$.

We will now derive a chain interpretation of the above results that will prove very useful later when we have to deal with the unknown correlation case. For the optimum filter coefficients $\{A_n^{(p)}\}_1^p$ and $\{B_n^{(p)}\}_1^p$, define the p-th order forward and backward residuals (see (5) and (9)) as the outputs of the forward and backward predictive error filters:

$$Y_k^{(p)} = -\sum_{n=0}^{p} A_n^{(p)} X_{k-n} = X_k - A_1^{(p)} X_{k-1} - \cdots - A_p^{(p)} X_{k-p} ,$$

$$Z_k^{(p)} = -\sum_{n=0}^{p} B_n^{(p)} X_{k-p+n} = X_{k-p} - B_1^{(p)} X_{k-p+1} - \cdots - B_p^{(p)} X_k . \qquad (17)$$

Then using (15), we can express

$$\begin{aligned} Y_k^{(p)} &= X_k - \sum_{n=1}^{p-1} A_n^{(p)} X_{k-n} - A_p^{(p)} X_{k-p} \\ &= X_k - \sum_{n=1}^{p-1} \left(A_n^{(p-1)} - A_p^{(p)} B_{p-n}^{(p-1)} \right) X_{k-n} - A_p^{(p)} X_{k-p} \\ &= -\sum_{n=0}^{p-1} A_n^{(p-1)} X_{k-n} + A_p^{(p)} \sum_{n=1}^{p} B_{p-n}^{(p-1)} X_{k-n} \\ &= Y_k^{(p-1)} + A_p^{(p)} \sum_{j=0}^{p-1} B_j^{(p-1)} X_{k-p+j} = Y_k^{(p-1)} - A_p^{(p)} Z_{k-1}^{(p-1)} . \end{aligned} \qquad (18)$$

And similarly

$$\begin{aligned} Z_k^{(p)} &= X_{k-p} - \sum_{n=1}^{p-1} B_n^{(p)} X_{k-p+n} - B_p^{(p)} X_k \\ &= X_{k-p} - \sum_{n=1}^{p-1} \left(B_n^{(p-1)} - B_p^{(p)} A_{p-n}^{(p-1)} \right) X_{k-p+n} - B_p^{(p)} X_k \\ &= -\sum_{n=0}^{p-1} B_n^{(p-1)} X_{k-p+n} + B_p^{(p)} \sum_{n=1}^{p} A_{p-n}^{(p-1)} X_{k-p+n} \\ &= Z_{k-1}^{(p-1)} + B_p^{(p)} \sum_{j=0}^{p-1} A_j^{(p-1)} X_{k-j} = Z_{k-1}^{(p-1)} - B_p^{(p)} Y_k^{(p-1)} . \end{aligned} \qquad (19)$$

Thus p-th order residuals $Y_K^{(p)}$ and $Z_K^{(p)}$ are related to the (p-1)th order residuals simply through the coefficients $A_p^{(p)}$ and $B_p^{(p)}$. A block diagram of the relationships in (18) and (19) is given in figure 1, where $z^{-1}I$ denotes an M x M matrix filter of unit delay.

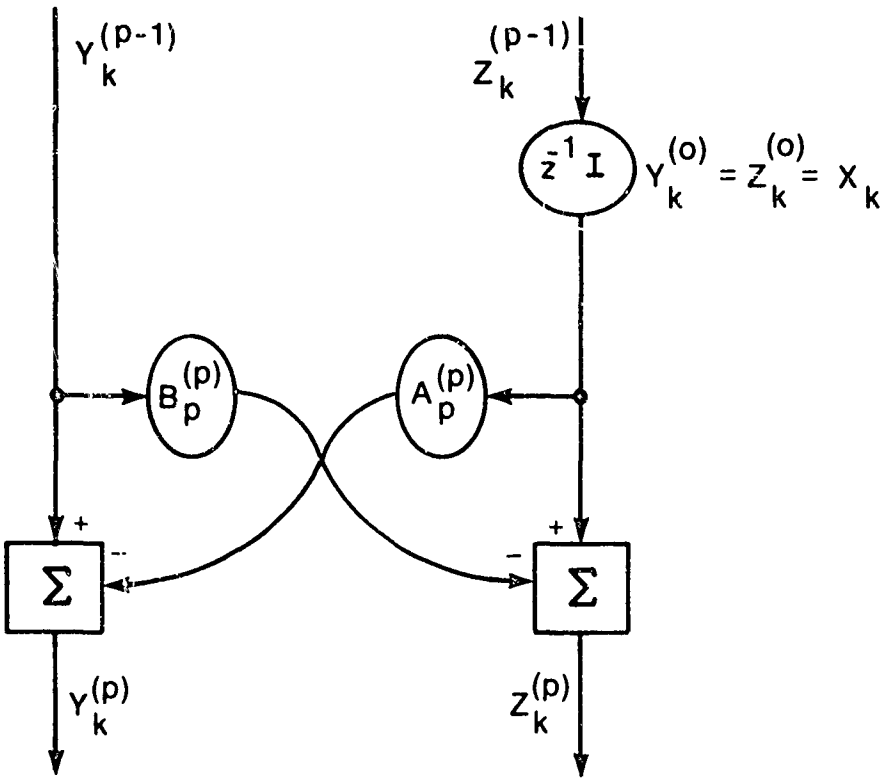

Figure 1. Chain Representation of Residuals

Thus matrix operators $A_p^{(p)}$ and $B_p^{(p)}$ can be interpreted as those coefficients which minimize

$$\overline{Y_k^{(p)H} Y_k^{(p)}} \quad \text{and} \quad \overline{Z_k^{(p)H} Z_k^{(p)}} \tag{20}$$

respectively, at the output of the p-th stage in figure 1, where $\{A_n^{(p)}\}_1^{p-1}$ and $\{B_n^{(p)}\}_1^{p-1}$ are determined by minimizations at lower order stages. $A_p^{(p)}$ and $B_p^{(p)}$ are called the partial correlation coefficients. Stated alternatively, stage by stage minimizations of (20), via choices of partial correlation coefficients $A_p^{(p)}$ and $B_p^{(p)}$, respectively, results in the same overall filter as if the powers in

$$-\sum_{n=0}^{p} A_n X_{k-n} \quad \text{and} \quad -\sum_{n=0}^{p} B_n X_{k-p+n} \tag{21}$$

were minimized by the choices of $\{A_n\}_1^p$ and $\{B_n\}_1^p$, respectively, each in one simultaneous optimization. This will furnish an important reference point

for the unknown correlation case in section 3.

If we let the transfer functions (z-transforms) to the outputs of the p-th stage in figure 1 be denoted by $\mathcal{H}_A^{(p)}(z)$ and $\mathcal{H}_B^{(p)}(z)$, it immediately follows, from figure 1 or equations (18) and (19), that

$$\begin{aligned}
\mathcal{H}_A^{(p)}(z) &= \mathcal{H}_A^{(p-1)}(z) - z^{-1} A_p^{(p)} \mathcal{H}_B^{(p-1)}(z), \\
\mathcal{H}_B^{(p)}(z) &= z^{-1} \mathcal{H}_B^{(p-1)}(z) - B_p^{(p)} \mathcal{H}_A^{(p-1)}(z), \\
\mathcal{H}_A^{(0)}(z) &= \mathcal{H}_B^{(0)}(z) = I.
\end{aligned} \tag{22}$$

In closed form, these predictive error filter transfer functions are expressible as (see (17))

$$\begin{aligned}
\mathcal{H}_A^{(p)}(z) &= -\sum_{n=0}^{p} z^{-n} A_n^{(p)} = I - \sum_{n=1}^{p} z^{-n} A_n^{(p)}, \\
\mathcal{H}_B^{(p)}(z) &= -\sum_{n=0}^{p} z^{n-p} B_n^{(p)} = -\sum_{j=0}^{p} z^{-j} B_{p-j}^{(p)} \\
&= z^{-p}\left[I - \sum_{n=1}^{p} z^{n} B_n^{(p)}\right].
\end{aligned} \tag{23}$$

2.2 PROPERTIES AND INTERPRETATIONS

Suppose that process $\{X_n\}$ were scaled according to

$$\tilde{X}_n = D X_n \tag{24}$$

where M x M matrix D is arbitrary, but invertible. Then the correlation of the scaled process is

$$\tilde{R}_k = \overline{\tilde{X}_n \tilde{X}_{n-k}^H} = D \overline{X_n X_{n-k}^H} D^H = D R_k D^H. \tag{25}$$

Now from (7), since the solutions $\{A_n^{(p)}\}$ and $\{\tilde{A}_n^{(p)}\}$ must satisfy

$$\sum_{n=1}^{p} \tilde{A}_n^{(p)} R_{m-n} = R_m, \quad 1 \le m \le p, \tag{26}$$

$$\sum_{n=1}^{p} \tilde{A}_n^{(p)} \tilde{R}_{m-n} = \tilde{R}_m \,, \quad 1 \le m \le p, \tag{27}$$

respectively, the solutions are related by a similarity transformation:

$$\tilde{A}_n^{(p)} = D A_n^{(p)} D^{-1}, \quad 1 \le n \le p. \tag{28}$$

This is called the scaling property. A similar property holds for the backward coefficients $\{B_n^{(p)}\}$.

An immediate by-product of the scaling property is that $A_n^{(p)}$ and $\tilde{A}_n^{(p)}$ have the same eigenvalues:

$$\det\left(\tilde{A}_n^{(p)} - \lambda I\right) = \det\left(D A_n^{(p)} D^{-1} - \lambda I\right) = \det\left(A_n^{(p)} - \lambda I\right). \tag{29}$$

Similarly, $B_n^{(p)}$ and $\tilde{B}_n^{(p)}$ have the same eigenvalues, regardless of scaling matrix D.

The remainder of this subsection will deal with the quantities U_p and V_p defined in (12), and C_p and D_p defined in (13). The quantity U_p can be interpreted physically as the correlation matrix of the p-th order forward residual; see (12), (5), and (17). Similarly, V_p is the correlation matrix of the p-th order backward residual; see (12), (9), and (17). That is,

$$U_p = \overline{Y_k^{(p)} Y_k^{(p)H}} \,, \quad V_p = \overline{Z_k^{(p)} Z_k^{(p)H}} \,. \tag{30}$$

Thus U_p and V_p are Hermitian:

$$U_p^H = U_p \,, \quad V_p^H = V_p \,; \tag{31}$$

and U_p and V_p are non-negative definite:

$$\eta^H U_p \eta = \eta^H \overline{Y_k^{(p)} Y_k^{(p)H}} \eta = \overline{\left|\eta^H Y_k^{(p)}\right|^2} \ge 0 \tag{32}$$

for any M x 1 matrix η. In appendix C, it is shown that simple recursions hold for U_p and V_p:

$$U_p = (I - A_p^{(p)} B_p^{(p)}) U_{p-1} \ , \ U_0 = R_0 \ ,$$
$$V_p = (I - B_p^{(p)} A_p^{(p)}) V_{p-1} \ , \ V_0 = R_0 \ . \tag{33}$$

It immediately follows from (33) that (see appendix C)

$$\det U_p = \det V_p \ , \ p \geq 0. \tag{34}$$

This property was proved in Ref. 8, page 240.

The quantity C_p defined in (13) can be interpreted as the cross-correlation matrix between the p-th order forward and backward residuals at one unit of delay:

$$\overline{Y_k^{(p)} Z_{k-1}^{(p)H}} = \sum_{n=0}^{p} \sum_{m=0}^{p} A_n^{(p)} \overline{X_{k-n} X_{k-1-p+m}^H} B_m^{(p)H}$$
$$= \sum_{m=0}^{p} \left(\sum_{n=0}^{p} A_n^{(p)} R_{p+1+m-n} \right) B_m^{(p)H} = - \sum_{n=0}^{p} A_n^{(p)} R_{p+1-n} = C_p \ , \tag{35}$$

where we have used (17), (2), (7), and (13). Similarly

$$\overline{Z_{k-1}^{(p)} Y_k^{(p)H}} = \sum_{m=0}^{p} \sum_{n=0}^{p} B_m^{(p)} \overline{X_{k-1-p+m} X_{k-n}^H} A_n^{(p)H}$$
$$= \sum_{n=0}^{p} \left(\sum_{m=0}^{p} B_m^{(p)} R_{p-1+n+m} \right) A_n^{(p)H} = - \sum_{m=0}^{p} B_m^{(p)} R_{-p-1+m} = D_p \ , \tag{36}$$

where we have used (17), (2), (11), and (13). It immediately follows from (35) and (36) that

$$D_p^H = C_p \ . \tag{37}$$

Thus it is not necessary to do the additional calculation of D_p in the solution given in (13).

Another interpretation of C_p is available as follows:

$$F_n^{(p)} = \overline{Y_k^{(p)} X_{k-m}^H} = - \sum_{n=0}^{p} A_n^{(p)} \overline{X_{k-n} X_{k-m}^H} = - \sum_{n=0}^{p} A_n^{(p)} R_{m-n} = \begin{cases} U_p \ , & m = 0 \\ 0 \ , & 1 \leq m \leq p \\ C_p \ , & m = p+1 \end{cases} \tag{38}$$

where we have employed (17), (2), (12), (7), and (13) in order. Thus the p-th order forward residual is uncorrelated with the p most recent past

values of the input, and the crosscorrelation at p + 1 units of delay is just C_p. Similarly, the backward residual satisfies

$$\overline{Z_m^{(p)} X_{k-p+m}^H} = -\sum_{n=0}^{p} B_n^{(p)} \overline{X_{k-p+n} X_{k-p+m}^H} = -\sum_{n=0}^{p} B_n^{(p)} R_{n-m} = \begin{cases} V_p, & m = 0 \\ 0, & 1 \le m \le p \\ D_p, & m = p+1 \end{cases} . \quad (39)$$

Yet another interpretation of C_p and D_p will be given in subsection 2.3.

As the order p in the linear prediction (4) increases (38) yields

$$F_m^{(p)} = \overline{Y_k^{(p)} X_{k-m}^H} \to \begin{cases} U_\infty, & m = 0 \\ 0, & 1 \le m \end{cases} \quad \text{as} \quad p \to \infty . \quad (40)$$

Therefore the autocorrelation matrix of the forward residual becomes

$$\overline{Y_k^{(p)} Y_{k-m}^{(p)H}} = -\sum_{n=0}^{p} \overline{Y_k^{(p)} X_{k-m-n}^H} A_n^{(p)H} \to \begin{cases} U_\infty, & m = 0 \\ 0, & 1 \le m \end{cases} \quad \text{as} \quad p \to \infty . \quad (41)$$

That is, p-th order residual $Y_k^{(p)}$ tends to white noise with a correlation matrix at zero time delay of value U_∞, which is not necessarily diagonal.

The Hermitian property in (37) allows us to combine (14) into the equation

$$A_p^{(p)} V_{p-1} = U_{p-1} B_p^{(p)H} , \quad (42)$$

where we utilized (31). This constraint on the partial correlation coefficients will be of paramount importance in the unknown correlation case. It immediately follows from (42) and (34) that

$$\det A_p^{(p)} = \det B_p^{(p)H} = \left(\det B_p^{(p)}\right)^* . \tag{43}$$

No such simple relation holds between $\det A_n^{(p)}$ and $\det B_n^{(p)}$ for $n<p$, except for $M = 1$, a univariate process.

2.3 EXTRAPOLATION OF CORRELATION VALUES

In subsection 2.1, we minimized the error in prediction (4) and found that for a p-th order prediction, knowledge of R_k for $|k| \le p$ was required; see (7). Now suppose that this is all the knowledge available about $\{R_k\}$; that is, suppose R_k is unknown for $|k|>p$. What can be done about approximating these unknown values?

One approach is as follows: we assume that the p-th order residual process $\{Y_k^{(p)}\}$ in (17) is white (i.e., uncorrelated for all non-zero delays), and that $A_p^{(p)} \neq 0$ (otherwise we could reduce the value of p). That is, we assume we can do nothing more in prediction by choosing more terms in the sum (4), which is tantamount to assuming maximum uncertainty (entropy) about the residual process $\{Y_k^{(p)}\}$. This is a very extensive assumption; we now investigate its ramifications.

We know from (38) that

$$F_m^{(p)} = \overline{Y_k^{(p)} X_{k-m}^H} = -\sum_{n=0}^{p} A_n^{(p)} R_{m-n}, \quad \text{all } m, \tag{44}$$

must satisfy

$$F_m^{(p)} = 0 \quad \text{for} \quad 1 \le m \le p. \tag{45}$$

Additionally, employing (17), the autocorrelation matrix of the p-th order residual is

$$\overline{Y_k^{(p)} Y_{k-j}^{(p)H}} = -\sum_{n=0}^{p} \overline{Y_k^{(p)} X_{k-j-n}^H} A_n^{(p)H} = -\sum_{n=0}^{p} F_{m+j}^{(p)} A_n^{(p)H}, \quad \text{all } j. \tag{46}$$

Now for j = 1, the white noise assumption on process $\{Y_k^{(p)}\}$ yields, via (46) and (45),

$$0 = -\sum_{n=0}^{p} F_{n+1}^{(p)} A_n^{(p)H} = -F_{p+1}^{(p)} A_p^{(p)H}; \quad \text{i.e.} \quad F_{p+1}^{(p)} = 0. \tag{47}$$

And for j = 2, the white assumption (in conjunction with (47)) yields

$$0 = -\sum_{n=0}^{p} F_{n+2}^{(p)} A_n^{(p)H} = -F_{p+2}^{(p)} A_p^{(p)H}; \quad \text{i.e.} \quad F_{p+2}^{(p)} = 0. \tag{48}$$

Continuing in this way, the white assumption is tantamount to assuming that

$$F_m^{(p)} = 0 \quad \text{for} \quad p+1 \le m. \tag{49}$$

Returning to expression (44), this means that we are assuming that

$$-\sum_{n=0}^{p} A_n^{(p)} R_{m-n} = 0 \quad \text{for} \quad p+1 \le m; \tag{50}$$

that is,

$$R_m = \sum_{n=1}^{p} A_n^{(p)} R_{m-n} \quad \text{for} \quad p+1 \le m. \tag{51}$$

Using more explicit notation, and denoting these assumed values of correlation as forward extrapolations $\{\hat{R}_m^{(p)}\}$, we have

$$\hat{R}_m^{(p)} \equiv \sum_{n=1}^{p} A_n^{(p)} \hat{R}_{m-n}^{(p)} \quad , \quad p+1 \le m, \tag{52}$$

where "starting values"

$$\hat{R}_m^{(p)} \equiv R_m \quad , \quad 0 \le m \le p. \tag{53}$$

Equation (52) is called the correlation recursion equation. It is interesting to note that the form of the correlation recursion (52) is identical to the form (4) for the individually predicted waveform values.

The correlation values in (52) are called the maximum entropy correlation extrapolations. The recursion is stable if and only if (see (23))

$$\det\left(I - \sum_{n=1}^{p} z^{-n} A_n^{(p)}\right) = \det \mathcal{A}_A^{(p)}(z) \tag{54}$$

possesses all its zeros within the unit circle in the complex z-plane; this property will be treated in subsection 2.4.

A similar procedure for backward correlation extrapolation, assuming that residual process $\{z_k^{(p)}\}$ is white, yields

$$\check{R}_{-m}^{(p)} \equiv \sum_{n=1}^{p} \check{B}_n^{(p)} \check{R}_{n-m}^{(p)} \quad , \quad p+1 \le m, \tag{55}$$

where

$$\check{R}_{-m}^{(p)} \equiv R_{-m} \quad , \quad 0 \le m \le p. \tag{56}$$

Backward recursion (55) is identical in form to the backward prediction (8). The recursion (55) is stable if and only if (see (23)).

$$\det\left(I - \sum_{n=1}^{p} z^{-n} B_n^{(p)}\right) = \det\left(z^{-p} a_B^{(p)}(z^{-1})\right) \tag{57}$$

possesses all its zeros within the unit circle.

As a special case of (52) and (53), the one-step forward extrapolated correlation based on a p-th order prediction is

$$\hat{R}_{p+1}^{(p)} = \sum_{n=1}^{p} A_n^{(p)} \hat{R}_{p+1-n}^{(p)} = \sum_{n=1}^{p} A_n^{(p)} R_{p+1-n} \ . \tag{58}$$

But from (13), we now can see that

$$C_p = -\sum_{n=0}^{p} A_n^{(p)} R_{p+1-n} = R_{p+1} - \hat{R}_{p+1}^{(p)} \ . \tag{59}$$

That is, C_p is the difference between the true correlation value R_{p+1} and the one-step forward extrapolated correlation $\hat{R}_{p+1}^{(p)}$ based upon knowledge of $\{R_k\}_{-p}^{p}$.

A similar procedure shows that

$$D_p = R_{-p-1} - \check{R}_{-p-1}^{(p)} \ . \tag{60}$$

That is, D_p is the difference between the true correlation value R_{-p-1} and the one-step backward extrapolated correlation $\check{R}_{-p-1}^{(p)}$ based upon knowledge of $\{R_k\}_{-p}^{p}$.

When (59) and (60) are combined with the Hermitian property in (37), we see that

$$\check{R}_{-p-1}^{(p)H} = \hat{R}_{p+1}^{(p)} \ . \tag{61}$$

This is a special case of the more general property (demonstrated in appendix D) that

$$\overleftarrow{R}_{-m}^{(p)H} = \hat{R}_{m}^{(p)} \quad , \quad p+1 \leq m \; ; \tag{62}$$

that is, the backward and forward extrapolated correlation matrices are Hermitians of each other. This is a desirable property of the extrapolations and is consistent with the same property, (2), which holds for the known correlation values, $\{R_k\}_{-p}^{p}$.

It was noted in (54) and (57) that the zeros of $\det \mathcal{A}_A^{(p)}(z)$ and $\det \mathcal{A}_B^{(p)}(z^{-1})$ must be within the unit circle in order that recursions (52) and (55), respectively, be stable. It is shown in appendix E that

$$\det\left(I - \sum_{n=1}^{p} z^{-n} A_n^{(p)}\right) = \det\left(I - \sum_{n=1}^{p} z^{-n} B_n^{(p)H}\right). \tag{63}$$

That we need consider only the zeros of one of these quantities; the location of these zeros is considered below.

It is also shown in appendix E that

$$\operatorname{tr} A_1^{(p)} = \left(\operatorname{tr} B_1^{(p)}\right)^* \tag{64}$$

and

$$\det A_p^{(p)} = \left(\det B_p^{(p)}\right)^* . \tag{65}$$

2.4 SPECTRAL APPROXIMATION

Equations (52) and (53) define the forward extrapolated correlations for all $m \geq 0$. We extend these to negative m via

$$\hat{R}_{m}^{(p)} = \hat{R}_{-m}^{(p)H} \quad , \quad m \leq 0, \tag{66}$$

which is consistent with (2). We will now use the Fourier transform of this infinite sequence, as in (3), as an approximation to the spectrum of process $\{X_n\}$. In appendix F, it is shown that the approximate spectrum is given by

$$G^{(p)}(f) = \Delta\, H_A^{(p)}(f)^{-1}\, U_p\, H_A^{(p)}(f)^{-1^H}, \quad |f| < \frac{1}{2\Delta}, \tag{67}$$

where

$$H_A^{(p)}(f) = -\sum_{n=0}^{p} \exp(-i2\pi f n\Delta) A_n^{(p)} \tag{68}$$

is the forward predictive error filter transfer function. Since U_p is nonnegative definite by (32), spectral approximation $G^{(p)}(f)$ is nonnegative definite for any f; it is also obviously Hermitian by (31). Thus the desirable properties of appendix A are achieved by approximation (67). In order to evaluate (67), one M x M matrix inverse (of $H_A^{(p)}(f)$) is needed at each value of f of interest.

A similar procedure applied to the backward correlation recursion of (55) and (56) yields the spectral approximation

$$G^{(p)}(f) = \Delta\, H_B^{(p)}(f)^{-1}\, V_p\, H_B^{(p)}(f)^{-1^H}, \quad |f| < \frac{1}{2\Delta}, \tag{69}$$

where

$$H_B^{(p)}(f) = -\sum_{n=0}^{p} \exp(-i2\pi f n\Delta) B_{p-n}^{(p)} \tag{70}$$

is the backward predictive error filter transfer function. Since the extrapolated correlations via (52) or (55) are equal, as shown in subsection 2.3, the same notation, $G^{(p)}(f)$, is used for both (67) and (69); however, we have two different factorizations for the unique spectral approximation $G^{(p)}(f)$.

In appendix F, it is also shown that the zeros of $\det \mathcal{H}_A^{(p)}(z)$ (see (22) and (23)) all lie inside the unit circle in the complex z-plane. Additionally, the poles of $\mathcal{H}_A^{(p)}(z)^{-1}$ all lie inside the unit circle, and the zeros of $\mathcal{H}_A^{(p)}(z)^{-1}$

all lie at z = 0. Thus the recursion (52) is stable. This point is discussed in Ref. 7, p. 132.

2.5 EXAMPLE

A simple example for M = 2 will be considered. Let the process be generated according to

$$X_k = G X_{k-1} + W_k , \tag{71}$$

where

$$G = \begin{bmatrix} .85 & -.75 \\ .65 & .55 \end{bmatrix} \tag{72}$$

and white noise W_k satisfies

$$\overline{W_k W_{k-m}^H} = \delta_{m0} I. \tag{73}$$

Then it may be shown that

$$R_m = G R_{m-1} + \delta_{m0} I, \quad m \geq 0, \tag{74}$$

with solution

$$R_0 = \begin{bmatrix} 25.135 & 4.862 \\ 4.862 & 21.643 \end{bmatrix}, \quad R_1 = \begin{bmatrix} 17.718 & -12.099 \\ 19.012 & 15.064 \end{bmatrix}. \tag{75}$$

By means of (7) and (11), we find

$$A_1^{(p)} = \begin{bmatrix} .85 & -.75 \\ .65 & .55 \end{bmatrix}, \quad B_1^{(p)} = \begin{bmatrix} .55930 & .75279 \\ -.64400 & .84070 \end{bmatrix}, \tag{76}$$

and $A_1^{(p)} = A_1^{(p)}$, $A_n^{(p)} = 0$, $2 \leq n \leq p$. We observe $A_1^{(p)} \neq B_1^{(p)}$, $A_1^{(p)} B_1^{(p)} \neq B_1^{(p)} A_1^{(p)}$, and $A_1^{(p)-1} \neq B_1^{(p)}$. The determinants of (76) are both .955.

Evaluation of (12) gives

$$U_1 = I, \quad V_1 = \begin{bmatrix} .91330 & .28934 \\ .28934 & 1.18659 \end{bmatrix}. \tag{77}$$

These matrices and their traces and eigenvalues are unequal, but their determinants are both 1.

3. UNKNOWN CORRELATION

In this section, the correlation values $\{R_k\}$ are unknown, and the only information available about the random process is a finite set of N data points $X_1, X_2, \ldots, X_N$, from which we have removed the sample mean. From these N data points, we desire an estimate of the spectrum G(f). But we cannot minimize or utilize any ensemble averages as was done in section 2, since we have only a finite segment of one member function to work with.

3.1 PHILOSOPHY OF APPROACH

For the known correlation case above, we had the set of normal equations

$$\left.\begin{aligned} \sum_{n=1}^{p} A_n^{(p)} R_{m-n} &= R_m \\ \sum_{n=1}^{p} B_n^{(p)} R_{n-m} &= R_{-m} \end{aligned}\right\}, \quad 1 \le m \le p, \tag{78A, 78B}$$

where $\{A_n^{(p)}\}_1^p$ and $\{B_n^{(p)}\}_1^p$ were the unknowns. Now in the unknown correlation case, we make a change by assuming that $A_p^{(p)}$ and $B_p^{(p)}$ are known* (along with R_m for $|m| \le p-1$, from lower order solutions), and by letting R_p and R_{-p} be unknown. The equations in the unknowns are still linear, and the solution is given by

$$\left.\begin{aligned} A_n^{(p)} &= A_n^{(p-1)} - A_p^{(p)} B_{p-n}^{(p-1)} \\ B_n^{(p)} &= B_n^{(p-1)} - B_p^{(p)} A_{p-n}^{(p-1)} \end{aligned}\right\}, \quad 1 \le n \le p-1 \qquad (p \ge 2), \tag{79A, 79B}$$

*The manner of specifying $A_p^{(p)}$ and $B_p^{(p)}$ will be considered in subsection 3.4.

$$R_p = \sum_{n=1}^{p} A_n^{(p)} R_{p-n} \quad , \tag{80A}$$

$$R_{-p} = \sum_{n=1}^{p} B_n^{(p)} R_{n-p} \quad . \tag{80B}$$

(It must be noted that R_k in this section denotes an estimate of the true (unknown) correlation value; for notational convenience, no distinguishing symbol has been added to R_k to emphasize this distinction.) However, we shall insist that the correlation estimates (80) that we obtain at the p-th stage satisfy

$$R_p = R_{-p}^H \quad , \tag{81}$$

in keeping with property (2). Since equations (78) and (81) are identical to those encountered in the known correlation case, the mathematical definitions and interrelationships employed there can be applied here also. However, some of the properties and physical interpretations may be different, since we are now dealing with estimates, rather than true values.

To solve (78), we begin by defining

$$R_c = \frac{1}{N} \sum_{k=1}^{N} X_k X_k^H = R_0^H \quad . \tag{82}$$

Now consider p=1 in (78); we have

$$A_1^{(1)} R_0 = R_1 \quad , \qquad B_1^{(1)} R_0 = R_{-1} \quad . \tag{83}$$

Now if $A_1^{(1)}$ and $B_1^{(1)}$ are known, we can compute unknowns R_1 and R_{-1}. But by constraint (81), $A_1^{(1)}$ and $B_1^{(1)}$ must be chosen such that

$$A_1^{(1)} R_0 = R_0 B_1^{(1)H} \quad . \tag{84}$$

Thus when we select $A_1^{(1)}$ and $B_1^{(1)}$, constraint (84) must be kept in mind; that is,

$A_1^{(1)}$ and $B_1^{(1)}$ cannot be specified independently of each other.

At stage $p(\geq 2)$, if $A_p^{(p)}$ is known (and $\{R_k\}_{1-p}^{p-1}$ are known from earlier stages with property $R_k = R_{-k}^H$, $0 \leq k \leq p-1$), we could solve the linear equations (78A) for $\{A_n^{(p)}\}_1^p$ and R_p, according to (79A) and (80A), where the lower order quantities in (79) and (80) are available from earlier stages. Similarly if $B_p^{(p)}$ is known, we use (79B) and (80B) to solve (78B). However, by (81), we must constrain the selection of $A_p^{(p)}$ and $B_p^{(p)}$.

To see exactly what constraint (81) implies about the selection of $A_p^{(p)}$ and $B_p^{(p)}$, notice that, for $p \geq 2$, (and defining $B_0^{(p-1)} = -I$)

$$R_p = \sum_{n=1}^{p} A_n^{(p)} R_{p-n} = \sum_{n=1}^{p-1} \left(A_n^{(p-1)} - A_p^{(p)} B_{p-n}^{(p-1)}\right) R_{p-n} + A_p^{(p)} R_0$$

$$= \sum_{n=1}^{p-1} A_n^{(p-1)} R_{p-n} - A_p^{(p)} \sum_{n=1}^{p} B_{p-n}^{(p-1)} R_{p-n}$$

$$= \sum_{n=1}^{p-1} A_n^{(p-1)} R_{p-n} - A_p^{(p)} \sum_{j=0}^{p-1} B_j^{(p-1)} R_j , \tag{35}$$

where we have employed (80A) and (79A). Now define forward extrapolated correlation estimates based on order p-1 according to (see (52) and (53))

$$R_m^{(p-1)} \equiv \sum_{n=1}^{p-1} A_n^{(p-1)} R_{m-n}^{(p-1)} \quad \text{for} \quad m \geq p, \tag{86}$$

where

$$R_m^{(p-1)} \equiv R_m , \quad 0 \leq m \leq p-1. \tag{87}$$

Then, in particular, the one-step forward extrapolated correlation estimate based on order p-1 is

$$R_p^{(p-1)} = \sum_{n=1}^{p-1} A_n^{(p-1)} R_{p-n}^{(p-1)} = \sum_{n=1}^{p-1} A_n^{(p-1)} R_{p-n} . \tag{88}$$

Also define (see (12))

$$V_{p-1} = -\sum_{n=0}^{p-1} B_n^{(p-1)} R_n \ . \tag{89}$$

This quantity has the physical interpretation as the estimate of the correlation matrix of the (p-1)th order backward residual at zero time delay (see (30)); its properties are considered in subsection 3.3. Then by means of (88) and (89), (85) can be expressed as

$$R_p = R_p^{(p-1)} + A_p^{(p)} V_{p-1} \ . \tag{90}$$

(This equation is similar to a combination of (14) and (59) for the known correlation case.)

At the same time, by (80B) and (79B) (and defining $A_0^{(p-1)} = -I$),

$$\begin{aligned} R_{-p} &= \sum_{n=1}^{p} B_n^{(p)} R_{n-p} = \sum_{n=1}^{p-1} \left(B_n^{(p-1)} - B_p^{(p)} A_{p-n}^{(p-1)} \right) R_{n-p} + B_p^{(p)} R_0 \\ &= \sum_{n=1}^{p-1} B_n^{(p-1)} R_{n-p} - B_p^{(p)} \sum_{n=1}^{p} A_{p-n}^{(p-1)} R_{n-p} \\ &= \sum_{n=1}^{p} B_n^{(p-1)} R_{n-p} - B_p^{(p)} \sum_{j=0}^{p-1} A_j^{(p-1)} R_{-j} \ . \end{aligned} \tag{91}$$

Now define backward extrapolated correlation estimates based on order p-1 as (see (55) and (56))

$$R_{-m}^{(p-1)} = \sum_{n=1}^{p-1} B_n^{(p-1)} R_{n-m}^{(p-1)} \quad \text{for} \quad m \geq p, \tag{92}$$

where

$$R_{-m}^{(p-1)} = R_{-m} \ , \ 0 \leq m \leq p-1. \tag{93}$$

Then, in particular,

$$R_{-p}^{(p-1)} = \sum_{n=1}^{p-1} B_n^{(p-1)} R_{n-p}^{(p-1)} = \sum_{n=1}^{p-1} B_n^{(p-1)} R_{n-p} \quad . \tag{94}$$

Also define (see (12))

$$U_{p-1} = - \sum_{n=0}^{p-1} A_n^{(p-1)} R_{-n} \quad . \tag{95}$$

This quantity is an estimate of the correlation matrix of the (p-1)th order forward residual at zero time delay (see (30)). Then by means of (94) and (95), (91) can be expressed as

$$R_{-p} = R_{-p}^{(p-1)} + B_p^{(p)} U_{p-1} \quad . \tag{96}$$

(This equation is similar to a combination of (14) and (60) for the known correlation case.) But now it can be shown (see appendix G) that the extrapolated correlation estimates in (88) and (94) satisfy

$$R_{-p}^{(p-1)H} = R_p^{(p-1)} \quad . \tag{97}$$

Therefore, if (81) is to be satisfied, (90) and (96) in conjunction with (97) force

$$A_p^{(p)} V_{p-1} = U_{p-1}^H B_p^{(p)H} \quad . \tag{98}$$

(This reduces to (84) for p=1.) Thus the selection procedure of $A_p^{(p)}$ and $B_p^{(p)}$ at the p-th stage must be done according to (98), where V_{p-1} and U_{p-1} are quantities already available from the (p-1)th stage, according to (89) and (95). The precise selection procedure will be undertaken in subsection 3.4.

3.2 COMPARATIVE FEATURES

There are alternative techniques to the estimation of the correlation matrices and the spectral density matrix that could be considered. For example, the standard Yule-Walker technique (e.g., Ref 2, page 186) uses correlation estimates

$$R_p = \frac{1}{N} \sum_K X_k X_{k-p}^H \quad , \tag{99}$$

where the sum is over all nonzero summands, and then solves recursively for $\{A_n^{(p)}\}_1^p$ and $\{B_n^{(p)}\}_1^p$ via the method in subsection 2.1. This apriori decision on the form (99) of the correlation estimate gives poorer spectral estimates for M=1 (Refs. 2 and 3), and probably does so for M>1. The estimated correlation matrix $[R_{m-n}]_0^p$ is Hermitian, block Toeplitz, and nonnegative definite:

$$\left[V_0^H \cdots V_p^H\right]\left[R_{m-n}\right]\begin{bmatrix} V_0 \\ \vdots \\ V_p \end{bmatrix} = \sum_{n,m=0}^{p} V_n^H R_{m-n} V_m = \sum_{n,m=0}^{p} V_n^H \left(\frac{1}{N}\sum_K X_{k-n} X_{k-m}^H\right) V_m$$

$$= \frac{1}{N}\sum_K \left|\sum_{n=0}^{p} V_n^H X_{k-n}\right|^2 \geq 0 \quad \text{for any } \{V_n\}_0^p \quad , \tag{100}$$

where V_n is $M \times 1$. However the stability of the correlation recursion (52) is unknown to this author. The estimate (99) is unchanged by the addition of more stages, that is, larger values of p.

Another technique would be to minimize the prediction error

$$Y_k = \sum_{n=0}^{p} A_n X_{k-n} \quad , \quad p+1 \le k \le N \qquad (A_0 = -I) \tag{101}$$

over the available data points directly, by choice of $\{A_n\}_1^p$. We have the error matrix

$$\frac{1}{N-p} \sum_{k=p+1}^{N} Y_k Y_k^H = \sum_{n,m=0}^{p} A_n S_{nm} A_m^H \quad , \tag{102}$$

where

$$S_{mn} \equiv \frac{1}{N-p} \sum_{k=p+1}^{N} X_{k-n} X_{k-m}^H \quad , \quad 0 \le m, n \le p. \tag{103}$$

The optimum coefficients for minimum trace of the error matrix, (102), are solutions of

$$\sum_{n=1}^{p} A_n^{(p)} S_{mn} = S_{m0} \quad , \quad 1 \le m \le p. \tag{104}$$

Matrix $[S_{mn}]_1^p$ is not block Toeplitz, and a significant computer problem exists for M>1 when it is noted that solution of linear equations (104) must be done anew for each different value of p. This was a good technique for spectral estimation when M=1 (see Ref. 3); however, computer time was greater than for the Burg technique. Moreover, stability of the correlation recursion (52) is unlikely in view of the (occasionally unstable) results for M=1 in Ref. 3.

This technique could be extended to include backward prediction in addition to (101). However, the lack of the block Toeplitz property and lack of stability make it a very undesirable technique.

The technique suggested here (in subsection 3.1) lets the correlation estimate be yielded according to solution (80), once partial correlation coefficients $A_p^{(p)}$ and $B_p^{(p)}$ have been specified. And we shall see in subsection 3.4 that these latter quantities are determined according to a physically meaningful minimization problem. Stability of the correlation recursion (52) has not been proved; however, numerous examples have all yielded stable solutions. The estimate (80) is unchanged by the addition of more stages, that is, larger values of p. And it will be seen that the current technique reduces to Burg's algorithm (Ref. 4) for M=1. Thus the current technique appears to be very attractive among those techniques that employ an all-pole representation of the input process.

3.3 PROPERTIES AND INTERPRETATIONS

The quantities U_{p-1} and V_{p-1} were defined in (95) and (89) and were interpreted as estimates of the correlation matrices of the (p-1)th order forward and backward residuals, respectively, at zero time delay. It is shown in appendix H that they satisfy the recurrence relations

$$\begin{aligned} U_p &= \left(I - A_p^{(p)} B_p^{(p)}\right) U_{p-1} , \qquad U_0 = R_0 , \\ V_p &= \left(I - B_p^{(p)} A_p^{(p)}\right) V_{p-1} , \qquad V_0 = R_0 , \end{aligned} \tag{105}$$

just as for the known correlation case. It is also shown that

$$U_p^H = U_p , \quad V_p^H = V_p , \tag{106}$$

and

$$\det U_p = \det V_p . \tag{107}$$

However, we are not able to prove U_p or V_p nonnegative definite without specifying the method by which $A_p^{(p)}$ and $B_p^{(p)}$ are selected; no relations like (30) and (32) exist here.

By means of (106), the constraint (98) on selection of $A_p^{(p)}$ and $B_p^{(p)}$ takes the form (see 42))

$$A_p^{(p)} V_{p-1} = U_{p-1} B_p^{(p)H} \quad . \tag{108}$$

This will be used in the next subsection.

3.4 EVALUATION OF PARTIAL CORRELATION COEFFICIENTS

We recall from subsection 2.1 that, in the known correlation case, the partial correlation coefficients $A_p^{(p)}$ and $B_p^{(p)}$ minimized

$$\text{tr } \overline{Y_k^{(p)} Y_k^{(p)H}} \quad \text{and} \quad \text{tr } \overline{Z_k^{(p)} Z_k^{(p)H}} \quad , \tag{109}$$

respectively, when lower order stages had already been optimized. We extend this idea to the unknown correlation case as follows: let (as in (18) and (19))

$$Y_k^{(0)} \equiv X_k \; , \; Z_k^{(0)} \equiv X_k \; , \quad 1 \le k \le N, \tag{110}$$

and for $p \ge 1$, define errors (residuals)

$$\left.\begin{aligned} Y_k^{(p)} &\equiv Y_k^{(p-1)} - A_p^{(p)} Z_{k-1}^{(p-1)} \\ Z_k^{(p)} &\equiv Z_{k-1}^{(p-1)} - B_p^{(p)} Y_k^{(p-1)} \end{aligned}\right\} \; , \; p+1 \le k \le N \tag{111}$$

The block diagram for (111) is identical to that in figure 1 on page 7.

Define for $p \geq 1$, the error (residual) matrix over the available data points as

$$E_p \equiv \frac{1}{N-p} \sum_{k=p+1}^{N} Y_k^{(p)} Y_k^{(p)H} = E_p^H ; \tag{112}$$

this nonnegative definite matrix is an unbiased estimator of $\overline{Y_k^{(p)} Y_k^{(p)H}}$.

Substitution of (111) in (112) yields

$$E_p = S_{p-1}^{(yy)} - A_p^{(p)} S_{p-1}^{(yz)H} - S_{p-1}^{(yz)} A_p^{(p)H} + A_p^{(p)} S_{p-1}^{(zz)} A_p^{(p)H} , \tag{113}$$

where

$$S_{p-1}^{(yy)} \equiv \frac{1}{N-p} \sum_{k=p+1}^{N} Y_k^{(p-1)} Y_k^{(p-1)H} = S_{p-1}^{(yy)H} , \tag{114A}$$

$$S_{p-1}^{(yz)} \equiv \frac{1}{N-p} \sum_{k=p+1}^{N} Y_k^{(p-1)} Z_{k-1}^{(p-1)H} , \tag{114B}$$

$$S_{p-1}^{(zz)} \equiv \frac{1}{N-p} \sum_{k=p+1}^{N} Z_{k-1}^{(p-1)} Z_{k-1}^{(p-1)H} = S_{p-1}^{(zz)H} . \tag{114C}$$

Also define for $p \geq 1$, error matrix

$$F_p \equiv \frac{1}{N-p} \sum_{k=p+1}^{N} Z_k^{(p)} Z_k^{(p)H} = F_p^H . \tag{115}$$

Substitution of (111) in (115) yields

$$F_p = S_{p-1}^{(zz)} - B_p^{(p)} S_{p-1}^{(yz)} - S_{p-1}^{(yz)H} B_p^{(p)H} + B_p^{(p)} S_{p-1}^{(yy)} B_p^{(p)H} . \tag{116}$$

Now error matrices E_p and F_p are Hermitian and nonnegative definite. Therefore matrix $\Lambda_a E_p \Lambda_a^H$ is Hermitian and nonnegative definite for any MxM weighting matrix Λ_a:

$$v^H (\Lambda_a E_p \Lambda_a^H) v = (\Lambda_a^H v)^H E_p (\Lambda_a^H v) \geq 0 \tag{117}$$

for any Mx1 matrix $\mathcal{V}$. Also since

$$tr(\Lambda_a E_p \Lambda_a^H) = tr(\Lambda_a^H \Lambda_a E_p) \equiv tr(\Lambda E_p), \tag{118}$$

only the product $\Lambda = \Lambda_a^H \Lambda_a$ matters in so far as the trace of $\Lambda_a E_p \Lambda_a^H$ is concerned; notice that Λ is Hermitian and nonnegative definite. We shall be interested in minimizing the traces of weighted error matrices $\Lambda_a E_p \Lambda_a^H$ and $\Gamma_a F_p \Gamma_a^H$; the exact choice of, and the reason for, weightings Λ_a and Γ_a will be undertaken in the next subsection.

Now if we were to minimize $tr(\Lambda_{p-1} E_p)$ by choice of $A_p^{(p)}$, we would find (see appendix B for method) that we must solve

$$\Lambda_{p-1} A_p^{(p)} S_{p-1}^{(zz)} = \Lambda_{p-1} S_{p-1}^{(yz)}, \tag{119}$$

and the choice of Λ_{p-1} would be irrelevant. Also, if we were to minimize $tr(\Gamma_{p-1} F_p)$ by choice of $B_p^{(p)}$, we would find that we must solve

$$\Gamma_{p-1} S_{p-1}^{(yy)} B_p^{(p)H} = \Gamma_{p-1} S_{p-1}^{(yz)}, \tag{120}$$

and the choice of Γ_{p-1} would be irrelevant. Furthermore, we would not satisfy constraint (108) generally. But since the behavior of error matrix F_p is just as important as that of E_p, we should take both matrices into account in any error minimization; in fact, for known correlation, recall that the determinants of residual matrices U_p and V_p were equal.

We therefore choose to minimize the sum of the traces of the weighted error matrices

$$\mathrm{tr}\left(\Lambda_{p-1} E_p\right) + \mathrm{tr}\left(\Gamma_{p-1} F_p\right) = \mathrm{tr}\left(\Lambda_{p-1} E_p + \Gamma_{p-1} F_p\right), \tag{121}$$

where Λ_{p-1} and Γ_{p-1} are Hermitian and nonnegative definite, by choice of $A_p^{(p)}$ and $B_p^{(p)}$ subject to constraint (108). If we let

$$A_p^{(p)} V_{p-1} = U_{p-1} B_p^{(p)H} \equiv G_p , \tag{122}$$

then we can express

$$\begin{aligned} &\Lambda_{p-1} E_p + \Gamma_{p-1} F_p = \\ &\Lambda_{p-1}\left[S_{p-1}^{(yy)} - G_p V_{p-1}^{-1} S_{p-1}^{(ya)H} - S_{p-1}^{(ya)} V_{p-1}^{-1} G_p^H + G_p V_{p-1}^{-1} S_{p-1}^{(aa)} V_{p-1}^{-1} G_p^H\right] \\ &+ \Gamma_{p-1}\left[S_{p-1}^{(ba)} - G_p^H U_{p-1}^{-1} S_{p-1}^{(by)} - S_{p-1}^{(ba)H} U_{p-1}^{-1} G_p + G_p^H U_{p-1}^{-1} S_{p-1}^{(bb)} U_{p-1}^{-1} G_p\right] \end{aligned} \tag{123}$$

in terms of the single unknown matrix G_p. Our problem therefore is to minimize the trace of (123) by choice of the single quantity G_p, subject to no constraints; we can then solve for the best coefficients according to

$$A_p^{(p)} = G_p V_{p-1}^{-1} , \quad B_p^{(p)} = G_p^H U_{p-1}^{-1} . \tag{124}$$

Also we can compute the correlation estimate from (90) and (88) according to

$$R_p = R_p^{(p-1)} + G_p . \tag{125}$$

In appendix I, it is shown that the minimum of the trace of (123) is realized when G_p is the solution of the bilinear matrix equation (Ref. 9)

$$G_p \alpha + \beta G_p = \mu + \nu , \tag{126}$$

where

$$\alpha = V_{p-1}^{-1} S_{p-1}^{(zz)} V_{p-1}^{-1} \Gamma_{p-1}^{-1}$$
$$\beta = \Lambda_{p-1}^{-1} U_{p-1}^{-1} S_{p-1}^{(yy)} U_{p-1}^{-1} \qquad (127)$$
$$\mu = S_{p-1}^{(yz)} V_{p-1}^{-1} \Gamma_{p-1}^{-1}$$
$$\nu = \Lambda_{p-1}^{-1} U_{p-1}^{-1} S_{p-1}^{(yz)} .$$

Uniqueness of the solution of (126) is considered in subsection 3.6. (It is interesting to note that the separate minimizations in (119) and (120) yield

$$G_p \alpha - \mu = \nu - \beta G_p = 0 . \qquad (128)$$

Thus whereas both these quantities had to be equal separately to the zero matrix, we now require only that they be equal to each other.)

For the special case of M=1 (a univariate process), (105) and (108) yield

$$U_p = V_p , \quad B_p^{(p)} = A_p^{(p)*} \qquad (M = 1). \qquad (129)$$

Then (126) and (127) can be solved for the scalar

$$G_p = \frac{(\Gamma_{p-1}^{-1} + \Lambda_{p-1}^{-1}) S_{p-1}^{(yz)}}{\Gamma_{p-1}^{-1} S_{p-1}^{(zz)} + \Lambda_{p-1}^{-1} S_{p-1}^{(yy)}} U_{p-1} \qquad (M = 1) . \qquad (130)$$

Now, if and only if

$$\Gamma_{p-1} = \Lambda_{p-1} \qquad (M = 1), \tag{131}$$

(130) reduces to Burg's algorithm (Ref. 4); in fact, it can be shown that (131) is the only choice of weights in (130) which guarantees a stable correlation recursion for M=1. Thus we shall insist that the weights satisfy (131) when we deal with their selection below.

3.5 WEIGHTING OF ERROR MATRICES

It is necessary to apply weighting to error matrices E_p and F_p in (112) and (115), prior to minimization of the trace in (121), for several reasons. First, without weighting, the larger amplitude components of errors (111) would receive most of the emphasis in the minimization; thus, some weighting inversely proportional to the component strengths is desired. Second, it is desired that stable correlation recursions result and that matrices U_p and V_p be nonnegative definite. Without weighting, it has been discovered (by an example to be presented in subsection 3.9) that both of these requirements can be violated. Third, we will insist that the scaling property introduced in subsection 2.2 hold for the unknown correlation case as well; that is, if

$$\tilde{X}_n = D X_n \ , \quad D \ \text{arbitrary} \ , \tag{132}$$

we shall insist that the coefficients satisfy

$$\left.\begin{aligned} \tilde{A}_n^{(p)} &= D A_n^{(p)} D^{-1} \\ \tilde{B}_n^{(p)} &= D B_n^{(p)} D^{-1} \end{aligned}\right\} \ , \ 1 \le n \le p, \quad \text{all } p. \tag{133}$$

The matrix equation (126) can be combined with (122) to yield the simultaneous set of equations

$$\begin{aligned} A_p^{(p)} S_{p-1}^{(ff)} V_{p-1}^{-1} \Gamma_{p-1}^{-1} + \Lambda_{p-1}^{-1} U_{p-1}^{-1} S_{p-1}^{(bb)} B_p^{(p)H} &= S_{p-1}^{(fb)} V_{p-1}^{-1} \Gamma_{p-1}^{-1} + \Lambda_{p-1}^{-1} U_{p-1}^{-1} S_{p-1}^{(bf)}, \\ A_p^{(p)} V_{p-1} - U_{p-1} B_p^{(p)H} &= 0. \end{aligned} \tag{134}$$

We now consider several possible choices of weightings Λ_{p-1} and Γ_{p-1} that tend to simplify the form of (134). The first choice is no weighting:

$$\Lambda_{p-1} = I, \quad \Gamma_{p-1} = I. \qquad \text{Choice 1} \tag{135}$$

The problem with this choice is that the weighting is not related to the error component strengths, and it may be readily verified that the solutions to (134) and (135) do not satisfy the scaling property (133). Also an unstable correlation recursion can occur. However, the solutions do reduce to Burg's algorithm for M=1; see (131).

Our next candidate weighting is

$$\Lambda_{p-1} = U_{p-1}^{-1}, \quad \Gamma_{p-1} = V_{p-1}^{-1}, \qquad \text{Choice 2} \tag{136}$$

which are Hermitian and are nonnegative definite if U_{p-1} and V_{p-1} are nonnegative definite. This weighting is inversely proportional to the component strengths, as desired; more will be said on this below. The equations (134) become

$$\begin{aligned} A_p^{(p)} S_{p-1}^{(bb)} + S_{p-1}^{(fb)} B_p^{(p)H} &= 2 S_{p-1}^{(fb)}, \\ A_p^{(p)} V_{p-1} - U_{p-1} B_p^{(p)H} &= 0. \end{aligned} \tag{137}$$

The solutions of (137) satisfy the scaling property (133), and they reduce to Burg's algorithm for M=1; (129) shows that (131) is satisfied for the choice (136). Although stability of the correlation recursions (52)

and (55), and nonnegative definiteness of U_p and V_p, have not been proven for general M≥2, no counter examples have been discovered.

We next consider

$$\Lambda_{p-1}^{-1} = S_{p-1}^{(yy)^{-1}} U_{p-1} \;, \quad \Gamma_{p-1}^{-1} = V_{p-1} S_{p-1}^{(zz)^{-1}} \;, \qquad \text{Choice 3} \tag{138}$$

in which case (134) becomes

$$\begin{aligned} A_p^{(p)} + B_p^{(p)H} &= S_{p-1}^{(yz)} S_{p-1}^{(zz)^{-1}} + S_{p-1}^{(yy)^{-1}} S_{p-1}^{(yz)} \;, \\ A_p^{(p)} V_{p-1} - U_{p-1} B_p^{(p)H} &= 0. \end{aligned} \tag{139}$$

However, the weighting (138) is not necessarily Hermitian, is not necessarily nonnegative definite, and is not directly related to the error component strengths. Also the solutions of (139) do not satisfy the scaling property. Furthermore, the solutions do not reduce to Burg's algorithm for M=1, and can yield unstable correlation recursions for M=1.

The last choice is

$$\Lambda_{p-1}^{-1} = U_{p-1} S_{p-1}^{(yy)^{-1}} U_{p-1} \;, \quad \Gamma_{p-1}^{-1} = V_{p-1} S_{p-1}^{(zz)^{-1}} V_{p-1} \;, \qquad \text{Choice 4} \tag{140}$$

which are Hermitian and nonnegative definite, and for which (134) becomes

$$\begin{aligned} A_p^{(p)} V_{p-1} + U_{p-1} B_p^{(p)H} &= S_{p-1}^{(yz)} S_{p-1}^{(zz)^{-1}} V_{p-1} + U_{p-1} S_{p-1}^{(yy)^{-1}} S_{p-1}^{(yz)} \;, \\ A_p^{(p)} V_{p-1} - U_{p-1} B_p^{(p)H} &= 0. \end{aligned} \tag{141}$$

This choice is a very interesting one in that the solutions of (141) are immediate and do not require that a bilinear matrix equation be solved. The weighting (140) is inversely proportional to the error component strengths, and the solutions of (141) do satisfy the scaling property. In fact, this choice is very close to Choice 2, since U_{p-1} and $S_{p-1}^{(yy)}$ are both estimates of the correlation matrix of process $\{Y_k^{(p-1)}\}$ at zero time delay, and should be

fairly close to each other. However, the solutions of (141) do not reduce to Burg's algorithm for M=1, and the correlation recursion (52) can be unstable, even for M=1. In fact, the solutions to (141) are identical to those for Choice 3 for M=1.

Therefore, of the four choices considered, only Choice 2 in (136) yields solutions that satisfy the scaling property (133) and reduces to Burg's algorithm for M=1. The stability of the correlation recursions has not been proved or disproved for choice (136) of weighting.

There is another strong reason for choosing weighting (136), which has to do with a whitening interpretation. We recall that U_{p-1} and V_{p-1}, defined in (95) and (89), are estimates of the correlation matrices of processes $\{Y_k^{(p-1)}\}$ and $\{Z_k^{(p-1)}\}$, respectively, at zero time delay. Now let (for non-negative definite U_{p-1} and V_{p-1})

$$U_{p-1} = \mathcal{U}_{p-1}\mathcal{U}_{p-1}^H \quad , \quad V_{p-1} = \mathcal{V}_{p-1}\mathcal{V}_{p-1}^H \quad , \tag{142}$$

where $\mathcal{U}_{p-1}$ and $\mathcal{V}_{p-1}$ are (lower triangular) square root matrices. Then scaled processes

$$\tilde{Y}_k^{(p-1)} \equiv \mathcal{U}_{p-1}^{-1} Y_k^{(p-1)} \quad , \quad \tilde{Z}_k^{(p-1)} \equiv \mathcal{V}_{p-1}^{-1} Z_k^{(p-1)} \quad , \quad p \le k \le N, \tag{143}$$

each have estimated correlation matrices at zero time delay equal to I; that is, all the components of $\{\tilde{Y}_k^{(p-1)}\}$ (or $\{\tilde{Z}_k^{(p-1)}\}$) have unit power and are uncorrelated with each other at zero time delay.

Now define, for $p+1 \le k \le N$,

$$q_k^{(p)} \equiv \mathcal{U}_{p-1}^{-1} Y_k^{(p)} = \mathcal{U}_{p-1}^{-1}\left(Y_k^{(p-1)} - A_p^{(p)} Z_{k-1}^{(p-1)}\right) = \tilde{Y}_k^{(p-1)} - \tilde{A}_p^{(p)} \tilde{Z}_{k-1}^{(p-1)} \quad , \tag{144}$$

where

$$\tilde{A}_p^{(p)} \equiv U_{p-1}^{-1} A_p^{(p)} V_{p-1} . \tag{145}$$

Also define the estimated correlation matrix at zero time delay of process $\{q_k^{(p)}\}$ as

$$\mathcal{E}_p \equiv \frac{1}{N-p} \sum_{k=p+1}^{N} q_k^{(p)} q_k^{(p)H} = \frac{1}{N-p} \sum_{k=p+1}^{N} U_{p-1}^{-1} Y_k^{(p)} Y_k^{(p)H} U_{p-1}^{-1H} = U_{p-1}^{-1} E_p U_{p-1}^{-1H}, \tag{146}$$

where we have used (144) and (112). Therefore

$$\mathrm{tr}\, \mathcal{E}_p = \mathrm{tr}\left(U_{p-1}^{-1} E_p U_{p-1}^{-1H}\right) = \mathrm{tr}\left(U_{p-1}^{-1H} U_{p-1}^{-1} E_p\right) = \mathrm{tr}\left(U_{p-1}^{-1} E_p\right), \tag{147}$$

where we have used (I-1) and (142). Thus, minimizing the trace of $U_{p-1}^{-1} E_p$, by choice of $A_p^{(p)}$, is equivalent to minimizing the trace of $\mathcal{E}_p$ by choice of $\tilde{A}_p^{(p)}$ (see (144)), where process $\{q_k^{(p)}\}$ is the error in prediction of (p-1)th order processes with estimated correlation matrices at zero time delay equal to I.

In a similar fashion, for $p+1 \le k \le N$,

$$Z_k^{(p)} \equiv V_{p-1}^{-1} Z_k^{(p)} = V_{p-1}^{-1} \left(Z_{k-1}^{(p-1)} - B_p^{(p)} Y_k^{(p-1)}\right) = \tilde{Z}_{k-1}^{(p-1)} - \tilde{B}_p^{(p)} \tilde{Y}_k^{(p-1)} \tag{148}$$

where

$$\tilde{B}_p^{(p)} \equiv V_{p-1}^{-1} B_p^{(p)} U_{p-1} . \tag{149}$$

And

$$J_p \equiv \frac{1}{N-p} \sum_{k=p+1}^{N} Z_k^{(p)} Z_k^{(p)H} = \frac{1}{N-p} \sum_{k=p+1}^{N} V_{p-1}^{-1} Z_k^{(p)} Z_k^{(p)H} V_{p-1}^{-1H} = V_{p-1}^{-1} F_p V_{p-1}^{-1} \tag{150}$$

with

$$\mathrm{tr}\, J_p = \mathrm{tr}\left(V_{p-1}^{-1} F_p\right). \tag{151}$$

If we solve (145) and (149) for $A_p^{(p)}$ and $B_p^{(p)}$, and then utilize constraint (108) along with (142), we find that the constraint takes the form

$$\tilde{A}_p^{(p)} = \tilde{B}_p^{(p)H} . \tag{152}$$

This could be used as the starting point in a minimization of error matrices $\mathcal{E}_p$ and $\mathcal{J}_p$. In fact, if we minimize the unweighted trace of $\mathcal{E}_p + \mathcal{J}_p$ by choice of $\tilde{A}_p^{(p)}$, we find the optimum choice to be given by

$$\tilde{A}_p^{(p)} \tilde{S}_{p-1}^{(zz)} + \tilde{S}_{p-1}^{(yy)} \tilde{A}_p^{(p)} = 2 \tilde{S}_{p-1}^{(yz)} , \tag{153}$$

where the notation is an obvious modification of (114). By employing (145), (143), and (142), we can show that (153) is equivalent to (137), as it must be. (This alternative approach may be useful for proving the stability of the correlation recursion.)

3.6 SOLUTION OF BILINEAR MATRIX EQUATION

If we substitute definitions (127) into bilinear matrix equation (126), and premultiply by $\Lambda_{p-1}^{1/2}$ and postmultiply by $\Gamma_{p-1}^{1/2}$, we obtain the equation

$$\tilde{G}_p \tilde{\alpha} + \tilde{\beta} \tilde{G}_p = \tilde{\mu} + \tilde{\nu} , \tag{154}$$

where

$$\begin{aligned}
\tilde{G}_p &\equiv \Lambda_{p-1}^{1/2} G_p \Gamma_{p-1}^{1/2} \\
\tilde{\alpha} &= \Gamma_{p-1}^{-1/2} V_{p-1}^{-1} S_{p-1}^{(xx)} V_{p-1}^{-1} \Gamma_{p-1}^{-1/2} \\
\tilde{\beta} &= \Lambda_{p-1}^{-1/2} U_{p-1}^{-1} S_{p-1}^{(ww)} U_{p-1}^{-1} \Lambda_{p-1}^{-1/2} \\
\tilde{\mu} &= \Lambda_{p-1}^{1/2} S_{p-1}^{(yx)} V_{p-1}^{-1} \Gamma_{p-1}^{-1/2} \\
\tilde{\nu} &= \Lambda_{p-1}^{-1/2} U_{p-1}^{-1} S_{p-1}^{(yw)} \Gamma_{p-1}^{1/2} .
\end{aligned} \tag{155}$$

Now the Hermitian matrices $\tilde{\alpha}$ and $\tilde{\beta}$ are non-negative definite; e.g.,

$$q^H \tilde{\alpha} q = \left(V_{p-1}^{-1} \Gamma_{p-1}^{-1/2} q\right)^H S_{p-1}^{(ss)} \left(V_{p-1}^{-1} \Gamma_{p-1}^{-1/2} q\right) \geq 0 \tag{156}$$

for any MX1 matrix q, since $S_{p-1}^{(ss)}$ is non-negative definite. We have employed the Hermitian property of V_{p-1} and Γ_{p-1} above; see (118) et seq. This means that the eigenvalues of $\tilde{\alpha}$ and $\tilde{\beta}$ must be non-negative. Therefore the solution of (154) exists and is unique (Ref. 10, eq. 3).

Solution of the bilinear matrix equation (126) or (154) has been addressed by many authors (Refs 9 - 17). In particular, for the equation involving MxM matrices,

$$XB + AX = C, \tag{157}$$

one form for the solution is given by

$$X = PQ^{-1}, \tag{158}$$

where

$$P = \sum_{k=0}^{M-1} (-1)^k A_k C B^{M-1-k},$$
$$Q = \sum_{k=0}^{M} (-1)^k a_k B^{M-k}, \tag{159}$$

are MxM matrices. The constants $\{a_k\}$ are given by (Ref. 18, pp. 87-88)

$$a_k = -\frac{1}{k} tr\left(A A_{k-1}\right), \quad 1 \leq k \leq M \qquad (a_0 = 1), \tag{160}$$

and the matrices $\{A_k\}$ are given by

$$A_k = A A_{k-1} + a_k I, \quad 1 \leq k \leq M \qquad (A_0 = I). \tag{161}$$

Here, M-2 full matrix multiplications are necessary when we note that $A_M = 0$

by the Cayley-Hamilton theorem.

For M = 2, (159) takes the form

$$\left.\begin{aligned} P &= CB - (A - \mathrm{tr}(A)\, I)\, C \\ Q &= (\mathrm{tr}\, A + \mathrm{tr}\, B) B + (\det A - \det B)\, I \end{aligned}\right\} \quad \text{for } M = 2, \tag{162}$$

where we have used the Cayley-Hamilton theorem to express

$$B^2 = \mathrm{tr}(B)\, B - \det(B)\, I \quad \text{for } M = 2. \tag{163}$$

Equations (162) and (158) are the forms used in the FORTRAN program for M = 2.

3.7 SPECTRAL ESTIMATION

Having obtained correlation estimates $\{R_m\}_0^p$ by means of (82) and (80A), we now extrapolate these, as in subsection 2.3 (equations (52) and (66)), to yield

$$\begin{aligned} R_m &= \sum_{n=1}^{p} A_n^{(p)} R_{m-n}, \quad p+1 \le m, \\ R_m &= R_{-m}^H, \quad m < 0. \end{aligned} \tag{164}$$

This defines an infinite sequence $\{R_m\}_{-\infty}^{\infty}$ which is assumed stable; its Fourier transform will be taken as the spectral estimate of the process under consideration. In a manner identical to that given in appendix F, it is found that

$$G^{(p)}(f) = \Delta \sum_{m=-\infty}^{\infty} \exp(-i 2\pi f m \Delta) R_m = \Delta\, H_A^{(p)}(f)^{-1}\, U_p\, H_A^{(p)}(f)^{-1^H}, \quad |f| < \frac{1}{2\Delta}, \tag{165}$$

where U_p and $H_A^{(p)}(f)$ are given by (95) and (68), respectively. It follows that

$$\int_{-\frac{1}{2\Delta}}^{\frac{1}{2\Delta}} df\, G^{(p)}(f) = R_0 = \text{sample power (80)}. \tag{166}$$

Also, as in subsection 2.4, an alternative factorization is available as

$$G^{(p)}(f) = \Delta\, H_B^{(p)}(f)^{-1}\, V_p\, H_B^{(p)}(f)^{-1^H}, \quad |f| < \frac{1}{2\Delta}, \tag{167}$$

where V_p and $H_B^{(p)}(f)$ are given in (89) and (70). If U_p or V_p is non-negative definite, then $G^{(p)}(f)$ is non-negative definite, as desired for a spectral estimate. Since (165) and (167) are equal, we concentrate henceforth on form (165).

Since

$$H_A^{(p)}(f)^{-1} = \frac{\text{Adj } H_A^{(p)}(f)}{\det H_A^{(p)}(f)} , \tag{168}$$

(165) can be expressed as

$$G^{(p)}(f) = \Delta \left| \det H_A^{(p)}(f) \right|^{-2} \left[\text{Adj } H_A^{(p)}(f)\right] U_p \left[\text{Adj } H_A^{(p)}(f)\right]^H . \tag{169}$$

Since $G^{(p)}(f)$ is Hermitian, matrix $G^{(p)}(f)$ need be computed only on and above its main diagonal, at each frequency of interest. Efficient computation of $H_A^{(p)}(f)$ by means of an FFT is undertaken in appendix J. It is shown that we need to perform M^2 N_F-point FFTs of p+1 non-zero numbers, in order to evaluate $H_A^{(p)}(f)$ at N_F frequency cells in the frequency range $\left(-\frac{1}{2\Delta}, \frac{1}{2\Delta}\right)$.

Real Multivariate Process

The results above have been derived for a complex multivariate process $\{X_k\}$. For a real multivariate process, U_p is real and $\{A_n^{(p)}\}_0^p$ are real. Then

$$H_A^{(p)}(-f) = H_A^{(p)}(f)^* \quad \text{for a real process,} \tag{170}$$

and

$$G^{(p)}(-f) = \Delta H_A^{(p)}(f)^{*-1} U_p H_A^{(p)}(f)^{*-1H}$$

$$= \Delta H_A^{(p)}(f)^{-1*} U_p H_A^{(p)}(f)^{-1H*} = G^{(p)}(f)^* \quad \text{for a real process.} \tag{171}$$

Thus we need compute matrix $G^{(p)}(f)$ only for $f \geq 0$, for a real multivariate process.

In order to avoid complex matrix multiplications, we develop (169) more explicitly; let

$$Adj\ H_A^{(p)}(f) = R_A(f) + i\, I_A(f), \tag{172}$$

where $R_A(f)$ and $I_A(f)$ are real MxM matrices at each f. Then since U_p is real, $U_p^T = U_p$, and upon substituting (172) in (169), we find

$$G^{(p)}(f) = A \left| \det H_A^{(p)}(f) \right|^{-2} \left[R_A(f) U_p R_A(f)^T + I_A(f) U_p I_A(f)^T + i\, M(f) - i\, M(f)^T \right] \tag{173}$$

for a real process,

where

$$M(f) \equiv I_A(f)\, U_p\, R_A(f)^T \tag{174}$$

Since M(f) is real, the quantity $i\,M(f) - i\,M(f)^T$ is zero on the main diagonal; therefore we need not compute the main diagonal of M(f). All the matrix multiplications in (173) are real.

Real Bivariate Process

We now further specialize to M = 2, a bivariate process. Let the real and imaginary parts of the filter transfer function $H_A^{(p)}$ be denoted by XX and YY, respectively (where these symbols are unrelated to X and Y introduced earlier); that is

$$H_A^{(p)}(f) \equiv XX(f) + i\, YY(f). \tag{175}$$

Then from (172), for 2 X 2 matrices,

$$R_A(f) = Re\left\{ Adj\, H_A^{(p)}(f) \right\} = Adj\ Re\left\{ H_A^{(p)}(f) \right\} = Adj\ XX(f) \equiv XX_A(f), \tag{176}$$

and

$$I_A(f) = \mathrm{Im}\left\{\mathrm{Adj}\, H_A^{(p)}(f)\right\} = \mathrm{Adj}\ \mathrm{Im}\left\{H_A^{(p)}(f)\right\} = \mathrm{Adj}\ YY(f) \equiv YY_A(f). \tag{177}$$

Substitution of (176) and (177) in (173) yields spectral estimate

$$G^{(p)}(f) = 4\left|\det H_A^{(p)}(f)\right|^{-2}\left[XX_A(f)U_p\,XX_A(f)^T + YY_A(f)U_p\,YY_A(f)^T + i\,M(f) - i\,M(f)^T\right]$$
$$\text{for a real bivariate process,} \tag{178}$$

where

$$M(f) = YY_A(f)\ U_p\ XX_A(f)^T. \tag{179}$$

The 2X2 matrices involved in (178) are all real, and $XX_A(f)$ and $YY_A(f)$ are the adjoints of the real and imaginary parts of $H_A^{(p)}(f)$, respectively. The form (178) is used in the program for the spectral estimate of a real bivariate process.

3.8 TERMINATION PROCEDURE

For unknown correlation, the correct value of p to use in (79) and (80) is unknown. We adopt the Akaike information criterion (AIC) derived in Ref. 19, page 719:

$$AIC_p = N \ln \det U_p + 2M^2p$$
$$= N \ln \det V_p + 2M^2p, \tag{180}$$

where we have utilized (107); namely, we compute AIC_p for $p = 0, 1, \ldots, p_{max}$, and we use that value of p, p_{best}, for which AIC_p is a minimum. Selection of p_{max} is discussed below.

For purposes of updating U_p and V_p, we can combine (105), (106), and (122) to yield

$$U_p = U_{p-1} - A_p^{(p)} G_p^H \,, \quad V_p = V_{p-1} - B_p^{(p)H} G_p \,, \tag{181}$$

in terms of the solution, G_p, of bilinear matrix equation (126).

At this point, it is worthwhile to review the procedure adopted here. From the actual data, we could have estimated the input correlation matrix via (99) (or some scaled version of it). Also we could have used (112) and (115) as error matrix estimates; in fact, these matrices are guaranteed Hermitian and non-negative definite. However, since $\det E_p \neq \det F_p$, we would have had to settle on some average like

$$\ln\left(\det E_p \cdot \det F_p\right)^{1/2} = \frac{1}{2}\left(\ln \det E_p + \ln \det F_p\right) \tag{182}$$

for purposes of the information criterion. As for the spectral estimate, we could have adopted, instead of (165), the quantity $\Delta H_A^{(p)}(f)^{-1} E_p \, H_A^{(p)}(f)^{-1H}$, or $\Delta H_B^{(p)}(f)^{-1} F_p \, H_B^{(p)}(f)^{-1H}$, for example.

Instead, we have chosen consistently to stick with the results of the normal equations (78). Thus the estimate of the input correlation matrix is obtained from (80)(and (82)); the estimates of the correlation matrices of the residuals are given by (89) and (95) (or more computationally convenient via (181)); and the spectral estimate is given in terms of U_p or V_p by (165) or (167), respectively, for $p=p_{best}$. The major gap in this procedure is that we have not proved that U_p or V_p is non-negative definite for Choice 2 of weighting in (136); however, no counter examples have been discovered.

Our selection of P_{max} is accomplished as follows: in ref. 1, page 575, Akaike is quoted as suggesting $P_{max} \leq 3N^{1/2}$ for M = 1, a univariate process. Since the number of coefficients evaluated is p, and the number of available data

points is N, this ratio was upper bounded by $3N^{-1/2}$. We extend this idea directly to the multivariate case: the number of scalar coefficients evaluated is $M^2 p$, and the number of available scalar data points is MN. Upper bounding this ratio by $3N^{-1/2}$, we find we should choose the filter order

$$p_{max} \leq \frac{3N^{1/2}}{M} \tag{183}$$

in terms of the number of data points, N, and the dimensionality of the time series, M.

3.9 EXAMPLES

It is worthwhile to summarize here the sequence of calculations required. For data $X_1, X_2, \ldots, X_N$ available (with the sample mean removed), we have

$$\begin{aligned}
&Y_k^{(a)} = Z_k^{(a)} = X_k, \quad 1 \leq k \leq N \\
&S_0^{(yy)} = \frac{1}{N-1}\sum_{k=2}^{N} X_k X_k^H = S_0^{(yy)H} \\
&S_0^{(zz)} = \frac{1}{N-1}\sum_{k=2}^{N} X_{k-1} X_{k-1}^H = \frac{1}{N-1}\sum_{k=1}^{N-1} X_k X_k^H = S_0^{(zz)H} \\
&S_0^{(yz)} = \frac{1}{N-1}\sum_{k=2}^{N} X_k X_{k-1}^H \\
&U_0 = V_0 = R_0 = \frac{1}{N}\sum_{k=1}^{N} X_k X_k^H = U_0^H = V_0^H .
\end{aligned} \tag{184}$$

Then for $p \geq 1$ and choice (136) of weighting,

$$\begin{aligned}
&\alpha = V_{p-1}^{-1} S_{p-1}^{(zz)} \\
&\beta = S_{p-1}^{(yy)} U_{p-1}^{-1} \\
&\mu = \nu = S_{p-1}^{(yz)} \\
&G_p \text{ via (126)} \\
&A_p^{(p)} = G_p V_{p-1}^{-1}, \quad B_p^{(p)} = G_p^H U_{p-1}^{-1} \\
&U_p = U_{p-1} - A_p^{(p)} G_p^H \\
&V_p = V_{p-1} - B_p^{(p)} G_p \\
&AIC_p = N \ln \det U_p + 2M^2 p
\end{aligned} \tag{185}$$

$$\left.\begin{aligned} Y_k^{(p)} &= Y_k^{(p-1)} - A_p^{(p)} Z_{k-1}^{(p-1)} \\ Z_k^{(p)} &= Z_{k-1}^{(p-1)} - B_p^{(p)} Y_k^{(p-1)} \end{aligned}\right\}, \quad p+1 \le k \le N \tag{186}$$

$$\begin{aligned} S_p^{(yy)} &= \frac{1}{N-p-1} \sum_{k=p+2}^{N} Y_k^{(p)} Y_k^{(p)H} = S_p^{(yy)H} \\ S_p^{(zz)} &= \frac{1}{N-p-1} \sum_{k=p+2}^{N} Z_{k-1}^{(p)} Z_{k-1}^{(p)H} = \frac{1}{N-p-1} \sum_{k=p+1}^{N-1} Z_k^{(p)} Z_k^{(p)H} = S_p^{(zz)H} \\ S_p^{(yz)} &= \frac{1}{N-p-1} \sum_{k=p+2}^{N} Y_k^{(p)} Z_{k-1}^{(p)H} . \end{aligned} \tag{187}$$

For $p = p_{max}$, it is not necessary to compute (186) through (187). When the best value of p, p_{best}, is found from AIC_p, we can then compute the spectral estimate (165).

We now consider an example for M = 2, N = 4:

$$X_1 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad X_3 = \begin{bmatrix} -1 \\ -2 \end{bmatrix}, \quad X_4 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \tag{188}$$

$$U_0 = V_0 = R_0 = \begin{bmatrix} 10 & 15 \\ 15 & 25 \end{bmatrix}$$

Then for weighting (136), we find

$$G_1 = -\frac{1}{12}\begin{bmatrix} 11 & 18 \\ 18 & 28 \end{bmatrix}, \quad A_1^{(1)} = B_1^{(1)} = -\frac{1}{6}\begin{bmatrix} 1 & 3 \\ 6 & 2 \end{bmatrix} \tag{189}$$

The eigenvalues of $A_1^{(1)}$ are $(-3 \pm \sqrt{73})/12$, which are both bounded by 1 in magnitude, as they must be for the correlation recursion (164) for p = 1 to be stable. Also,

$$U_1 = V_1 = \frac{1}{72}\begin{bmatrix} 7 & 6 \\ 6 & 16 \end{bmatrix}, \tag{190}$$

which is non-negative definite. Thus, for weighting (136), all the desirable properties are realized.

However, for no weighting, (135), we find, for the same example (188),

$$G_1 = -\frac{1}{18}\begin{bmatrix} 20 & 30 \\ 30 & 43 \end{bmatrix}, \quad A_1^{(1)} = B_1^{(1)} = \frac{1}{9}\begin{bmatrix} -10 & 0 \\ -21 & 4 \end{bmatrix}. \tag{191}$$

The eigenvalues of $A_1^{(1)}$ are 4/9 and -10/9; since the latter is larger than 1 in magnitude, the recursion $R_m = A_1^{(1)} R_{m-1}$, $m \geq 1$, is unstable. Also

$$U_1 = V_1 = -\frac{1}{162}\begin{bmatrix} 38 & 57 \\ 57 & 53 \end{bmatrix}, \tag{192}$$

which is not a non-negative definite matrix. It is found that the spectral estimate obtained from (165) has frequency ranges where the two autospectra (diagonal terms of (165)) are negative, and where the magnitude-squared coherence can be negative or greater than 1. These are all unacceptable.

For the alternative example for M = 2, N = 4, of

$$X_1 = \begin{bmatrix} -.25 \\ -1.19 \end{bmatrix}, \quad X_2 = \begin{bmatrix} -1.25 \\ 81 \end{bmatrix}, \quad X_3 = \begin{bmatrix} 75 \\ 81 \end{bmatrix}, \quad X_4 = \begin{bmatrix} .75 \\ -43 \end{bmatrix}, \tag{193}$$

and no weighting, we find a stable correlation recursion, but U_1 and V_1 are not non-negative definite, and values of the magnitude-squared coherence greater than 1 are realized in some frequency ranges. Because of these unacceptable behaviors, the choice of no weighting, (135), is discarded from future consideration.

An example for M = 2, N = 100, and weighting (136), generated via (71) - (73) of subsection 2.5 yielded the results below; the program and its output are given in appendix K. We find p_{best}=1 and

$$A_1^{(1)} = \begin{bmatrix} .87151 & -.77024 \\ .63432 & .56035 \end{bmatrix}, \quad B_1^{(1)} = \begin{bmatrix} .56613 & .77098 \\ -.63142 & .86573 \end{bmatrix}, \qquad (194)$$

$$U_1 = .09110 \begin{bmatrix} 97698 & -00867 \\ -00867 & 1.02364 \end{bmatrix}, \quad V_1 = .09110 \begin{bmatrix} .93178 & .34488 \\ 34488 & 1.20087 \end{bmatrix}. \qquad (195)$$

It is worthwhile to compare these estimates for N = 100 with the exact values in (76) and (77). The scale factor .09110 in (195) is unimportant and is due to the fact that the white noise used here had variance 1/12 rather than 1 as in (73); except for the scale factor, the matrices in (195) have determinants equal to 1. The estimated magnitude-squared coherence reaches a maximum of .999745, versus the true peak of .999013.

Observations from other examples of real bivariate processes have pointed out that: the eigenvalues of $A_1^{(n)}$ and $B_1^{(n)}$ are identical and are bounded by 1 in magnitude; the eigenvalues of $A_p^{(n)}$ and $B_p^{(n)}$ are not identical for $p \geq 2$, and can be larger than 1 in magnitude; and the eigenvalues of $A_n^{(n)}$ and $B_n^{(n)}$ for $n<p$ can be larger than 1 in magnitude.

Timing Results

Some sample execution times on a UNIVAC 1108 for SUBROUTINE PCC, which evaluates the partial correlation coefficients, are presented below for M = 2, a bivariate real process.

Table 1. Timing of Subroutine PCC

N	p_{max}	Time of Execution (sec)
100	10	0.25
100	15	0.35
1000	10	2.63
1000	40	9.23
10000	50	120
10000	150	326

The execution time is almost linearly proportional to N and p_{max}. The execution time for PEFTF was 1.25 seconds, and that for SDM was 0.55 seconds, both for N_F = 1024 frequency cells; see appendix K for program.

4. SUMMARY

A method for multivariate linear predictive spectral analysis, employing weighted forward and backward averaging, has been presented and programmed in FORTRAN. The method constitutes a generalization of Burg's univariate algorithm (Ref. 4) to the multivariate case.

The choice of weighting in the error minimization is very important, and several candidates have been considered. The weighting retained, (136), is the only one of those considered that satisfies both the scaling property (133) for all M, and reduces to Burg's algorithm for M = 1. Also, the weighting retained is equivalent to minimizing the unweighted traces of

error processes that are the differences of approximately white processes; in fact, (136) could be used as the starting point of the error minimization.

The major gaps in the analysis are that we have not proved that U_p and V_p are non-negative definite, and we have not proved that correlation recursion (164) is stable; however, no counterexamples have been encountered. The major analytical block in this endeavor is the bilinear matrix equation, (126), which requires special treatment for its solution.

Appendix A

PROPERTIES OF A SPECTRAL DENSITY MATRIX

Suppose an arbitrary linear filter with impulse response $\{H_n\}$ is excited by input $\{X_k\}$. The output at time $k\Delta$ is

$$Y_k = \sum_n H_n X_{k-n} , \tag{A-1}$$

where the sum is over all non-zero summands. X_k and Y_k are M x 1 matrices, whereas H_n is M x M. In steady state, the spectra of the processes in (A-1) are related by

$$G_Y(f) = H(f)\, G_X(f)\, H(f)^H , \tag{A-2}$$

where transfer function

$$H(f) \equiv \sum_n \exp(-i2\pi f n\Delta)\, H_n , \tag{A-3}$$

and f frequency in Hz and is real.

Now

$$G_X(f)^H = \Delta \sum_{k=-\infty}^{\infty} \exp(i2\pi f k\Delta) R_k^H = \Delta \sum_{n=-\infty}^{\infty} \exp(-i2\pi f n\Delta) R_n = G_X(f), \tag{A-4}$$

where we have employed (2). Thus $G_X(f)$ is Hermitian at any value of f. Similarly $G_Y(f)$ is Hermitian at any f.

Also

$$R_0^{(Y)} \equiv \overline{Y_k Y_k^H} \tag{A-5}$$

is non-negative definite for any H(f), because

$$q^H R_0^{(Y)} q = q^H \overline{Y_k Y_k^H}\, q = \overline{\left| q^H Y_k \right|^2} \geq 0 \tag{A-6}$$

for any M x 1 column matrix q . Therefore

$$R_0^{(Y)} = \int_{-\frac{1}{2\Delta}}^{\frac{1}{2\Delta}} df\, G_Y(f) = \int_{-\frac{1}{2\Delta}}^{\frac{1}{2\Delta}} df\, H(f) G_x(f) H(f)^H \qquad \text{(A-7)}$$

is non-negative definite for any H(f). It then follows that

$$G_x(f) \text{ is non-negative definite for all } f. \qquad \text{(A-8)}$$

To prove this, assume that $G_x(f_1)$ is not non-negative definite; than if we choose $H(f) \sim I\, \delta(f-f_1)$, that is, an impulsive transfer function near frequency f_1, we get $R_0^{(Y)} \sim G_x(f_1)$ from (A-7), which contradicts the conclusion that $R_0^{(Y)}$ must be non-negative definite.

Thus a spectral density matrix must always be Hermitian and non-negative definite for all f. In particular, this implies that all the auto spectra (diagonal terms of the matrix) must be real and non-negative. It also implies that all coherences are bounded by unity in magnitude.

Appendix B

MINIMIZATION OF TRACE OF ERROR MATRIX

From (4) and (5), we have

$$Y_k = X_k - \sum_{m=1}^{P} A_m X_{k-m} = X_k - a\chi_k \tag{B-1}$$

where

$$a \equiv [A_1 \cdots A_p], \quad \chi_k \equiv \begin{bmatrix} X_{k-1} \\ \vdots \\ X_{k-p} \end{bmatrix}. \tag{B-2}$$

Let

$$\overline{X_k \chi_k^H} = \mathcal{C}, \quad \overline{\chi_k \chi_k^H} = Q. \tag{B-3}$$

Here, $\mathcal{C}$ is M x Mp, χ_k is Mpx1, $\mathcal{C}$ is MxMp, and Q is MpxMp. We notice that $Q^H = Q$, and $v^H Q v = \overline{|v^H \chi_k|^2} > 0$ for any Mpx1 matrix $v \neq 0$, if no exact linear relation exists between the elements of $X_{k-1}, \ldots, X_{k-p}$; that is, Q is Hermitian and positive definite.

Now

$$\overline{Y_k Y_k^H} = \overline{(X_k - a\chi_k)(X_k^H - \chi_k^H a^H)}$$

$$= R_0 - a\mathcal{C}^H - \mathcal{C}a^H + aQa^H \tag{B-4}$$

$$= R_0 - \mathcal{C}Q^{-1}\mathcal{C}^H + (a - \mathcal{C}Q^{-1})Q(a - \mathcal{C}Q^{-1})^H. \tag{B-5}$$

Let

$$a - \mathcal{C}Q^{-1} = \begin{bmatrix} v_1^H \\ \vdots \\ v_M^H \end{bmatrix} \tag{B-6}$$

where v_j is an Mpx1 matrix. Then for the M x M matrix in (B-5),

$$L \equiv (a - cQ^{-1})Q(a - cQ^{-1})^H = [l_{jm}]_{j,m=1}^{M}, \tag{B-7}$$

where complex scalar

$$l_{jm} = v_j^H Q v_m. \tag{B-8}$$

The real quantity $l_{jj} = v_j^H Q v_j > 0$ for any $v_j \neq 0$, since Q is Hermitian and positive definite; the minimum value of l_{jj} is zero and is attained if and only if $v_j = 0$. Therefore, tr L is minimized, attaining value zero, by the choice $v_j = 0, 1 \leq j \leq M$. Thus $\text{tr}\,\overline{Y_k Y_k^H} = \overline{Y_k^H Y_k}$ is minimized by the choice of a as

$$a_{opt} \equiv [A_1^{(p)} \cdots A_p^{(p)}] = cQ^{-1}, \tag{B-9}$$

since the leading two terms in (B-5) are independent of a.

Then we have opt L = 0 and

$$\text{opt}\ \overline{Y_k Y_k^H} = R_0 - cQ^{-1}c^H = R_0 - a_{opt}c^H = R_0 - a_{opt}Q a_{opt}^H. \tag{B-10}$$

Also

$$\min \overline{Y_k^H Y_k} = \text{tr opt}\ \overline{Y_k Y_k^H} = \text{tr}(R_0 - cQ^{-1}c^H) = \overline{X_k^H X_k} - \text{tr}(cQ^{-1}c^H) \tag{B-11}$$

It should be noted that the solution (B-9) is attainable directly from (B-4) if the coefficient of a^H (or a) is set equal to zero; this observation will be useful later.

Equations (B-9) and (B-10) can be developed as follows:

$a_{opt}Q = c$ yields, with the use of (B-2) and (B-3),

$$[A_1^{(p)} \cdots A_p^{(p)}] \begin{bmatrix} R_0 & R_1 & \cdots & R_{p-1} \\ R_{-1} & \ddots & & \\ \vdots & & & \\ R_{1-p} & & & R_0 \end{bmatrix} = [R_1 \cdots R_p]; \tag{B-12}$$

that is,

$$\sum_{n=1}^{p} A_n^{(p)} R_{m-n} = R_m , \quad 1 \le m \le p. \tag{B-13}$$

And (B-10) can be expressed as

$$\text{opt } \overline{Y_k Y_k^H} = R_0 - \left[A_1^{(p)} \cdots A_p^{(p)}\right] \begin{bmatrix} R_1^H \\ \vdots \\ R_p^H \end{bmatrix} = R_0 - \sum_{n=1}^{p} A_n^{(p)} R_n . \tag{B-14}$$

Equations (B-13) and (B-14) are the main results of this appendix.

If an exact linear relation exists between the elements of $X_{k-1}, \cdots, X_{k-p}$, then

$$X_{k-1} = \sum_{j=2}^{p} G_j X_{k-j} \quad \text{for some } \{G_j\}_2^p \neq 0. \tag{B-15}$$

In this case, (B-1) yields

$$Y_{k-1} = X_{k-1} - \sum_{n=1}^{p} A_n X_{k-1-n} = X_{k-1} - \sum_{j=2}^{p} A_{j-1} X_{k-j} - A_p X_{k-1-p} . \tag{B-16}$$

Therefore we can get zero error by choosing

$$A_n^{(p)} = \left\{ \begin{array}{l} G_{n+1} , \quad 1 \le n \le p-1 \\ 0, \quad n = p \end{array} \right\} . \tag{B-17}$$

Thus $A_p^{(p)} = 0$ if an exact linear relation exists between the elements of $X_{k-1}, \cdots, X_{k-p}$.

Also we have the following general theorem:

No exact linear relation between elements of $X_k, \cdots, X_{k-p}$ $\iff$ $\begin{bmatrix} R_0 & \cdots & R_p \\ \vdots & \searrow & \\ R_p & & R_0 \end{bmatrix}$ is positive definite. (B-18)

To prove this, let

$$\mathbf{X}_k = \begin{bmatrix} X_k \\ \vdots \\ X_{k-p} \end{bmatrix} , \quad D = \begin{bmatrix} d_1 \\ \vdots \\ d_{m(p+1)} \end{bmatrix} \tag{B-19}$$

Then $F_K \equiv D^H X_K$ is a scalar. Now if and only if an exact linear relation exists, $F_K = 0$ for some $D \neq 0$, no matter which member function of the ensemble we select (with probability one). We also notice that

$$\overline{|F_K|^2} = D^H \overline{X_K X_K^H} D \tag{B-20}$$

and that the ensemble average in (B-20) is equal to the matrix in (B-18).

Assume that $F_K \neq 0$ for any $D \neq 0$. Then $\overline{|F_K|^2} > 0$ for any $D \neq 0$, and the right-hand side of (B-20) is positive for any $D \neq 0$. Therefore $\overline{X_K X_K^H}$ is positive definite.

Conversely if $\overline{X_K X_K^H}$ is positive definite, the right-hand side of (B-20) is positive for any $D \neq 0$. Then $\overline{|F_K|^2} > 0$ for any $D \neq 0$, yielding $F_K \neq 0$ for any $D \neq 0$.

Appendix C

INTERRELATIONSHIPS OF U_p AND V_p

We start with the definition (12) and develop U_p as

$$
\begin{aligned}
U_p &= R_0 - \sum_{m=1}^{p} A_m^{(p)} R_{-m} \\
&= R_0 - \sum_{m=1}^{p-1} \left(A_m^{(p-1)} - A_p^{(p)} B_{p-m}^{(p-1)}\right) R_{-m} - A_p^{(p)} R_{-p} \qquad \text{(by (15))} \\
&= R_0 - \sum_{m=1}^{p-1} A_m^{(p-1)} R_{-m} + A_p^{(p)} \sum_{m=1}^{p} B_{p-m}^{(p-1)} R_{-m} \\
&= U_{p-1} + A_p^{(p)} \sum_{j=0}^{p-1} B_j^{(p-1)} R_{j-p} \qquad \text{(by (12))} \\
&= U_{p-1} - A_p^{(p)} D_{p-1} \qquad \text{(by (13))} \\
&= U_{p-1} - A_p^{(p)} B_p^{(p)} U_{p-1} \qquad \text{(by (14))} \\
&= \left(I - A_p^{(p)} B_p^{(p)}\right) U_{p-1} \qquad \text{(C-1)}
\end{aligned}
$$

This relation holds for $p \geq 1$, with $U_0 = R_0$. A similar derivation for V_p yields

$$V_p = \left(I - B_p^{(p)} A_p^{(p)}\right) V_{p-1} \quad , \; p \geq 1 \; ; \quad V_0 = R_0 . \qquad \text{(C-2)}$$

The determinant of U_p is given by

$$
\begin{aligned}
\det U_p &= \det\left(I - A_p^{(p)} B_p^{(p)}\right) \det U_{p-1} \\
&= \det A_p^{(p)} \; \det\left(A_p^{(p)-1} - B_p^{(p)}\right) \cdot \det U_{p-1} \; , \qquad \text{(C-3)}
\end{aligned}
$$

whereas the determinant of V_p is

$$
\begin{aligned}
\det V_p &= \det\left(I - B_p^{(p)} A_p^{(p)}\right) \cdot \det V_{p-1} \\
&= \det\left(A_p^{(p)-1} - B_p^{(p)}\right) \det A_p^{(p)} \cdot \det V_{p-1} \qquad \text{(C-4)}
\end{aligned}
$$

Now if det U_{p-1} = det V_{p-1}, then (C-3) and (C-4) indicate that

$$\det U_p = \det V_p . \qquad \text{(C-5)}$$

But since $U_o = V_o = R_o$, det U_o = det V_o. Therefore (C-5) holds for $p \geq 0$, by induction.

Appendix D

HERMITIAN PROPERTY OF EXTRAPOLATED CORRELATIONS

We know that

$$R_{-k}^{H} = R_k \quad \text{for} \quad |k| \le p. \tag{D-1}$$

We then solve

$$\sum_{n=1}^{p} A_n^{(p)} R_{k-n} = R_k \;, \quad 1 \le k \le p \tag{D-2}$$

for $\{A_n^{(p)}\}_1^p$, and set

$$\hat{R}_k^{(p)} = \sum_{n=1}^{p} A_n^{(p)} \hat{R}_{k-n}^{(p)} \quad \text{for all } k \ge 1\,; \; \hat{R}_k^{(p)} \equiv R_k \text{ for } |k| \le p. \tag{D-3}$$

We then define

$$\hat{R}_{-k}^{(p)} = \hat{R}_k^{(p)\,H} \quad \text{for} \quad p+1 \le k. \tag{D-4}$$

In a similar fashion for the backward case, we solve

$$\sum_{n=1}^{p} B_n^{(p)} R_{n-k} = R_{-k} \;, \quad 1 \le k \le p \tag{D-5}$$

for $\{B_n^{(p)}\}_1^p$, and set

$$\check{R}_{-k}^{(p)} = \sum_{n=1}^{p} B_n^{(p)} \check{R}_{n-k}^{(f)} \quad \text{for all } k \ge 1\,; \; \check{R}_{-k}^{(p)} \equiv R_{-k} \text{ for } |k| \le p. \tag{D-6}$$

We then define

$$\check{R}_k^{(p)} = \check{R}_{-k}^{(p)\,H} \quad \text{for} \quad p+1 \le k. \tag{D-7}$$

We know from the definitions above that

$$\check{R}^{(p)H}_{-k} = \hat{R}^{(p)}_{k} \quad \text{for} \quad |k| \le p. \tag{D-8}$$

Now we assume that

$$\check{R}^{(p)H}_{-k} = \hat{R}^{(p)}_{k} \quad \text{for} \quad |k| \le m, \text{ where } m \ge p; \tag{D-9}$$

that is, from (D-6) and (D-3),

$$\check{R}^{(p)H}_{-k} = \sum_{n=1}^{p} \check{R}^{(p)H}_{n-k} B^{(p)H}_{n} = \sum_{n=1}^{p} A^{(p)}_{n} \hat{R}^{(p)}_{k-n} = \hat{R}^{(p)}_{k} \quad \text{for} \quad 1 \le k \le m. \tag{D-10}$$

Now from recursion definition (D-6),

$$\begin{aligned}
\check{R}^{(p)H}_{-m-1} &= \sum_{n=1}^{p} \check{R}^{(p)H}_{n-m-1} B^{(p)H}_{n} \\
&= \sum_{n=1}^{p} \hat{R}^{(p)}_{m+1-n} B^{(p)H}_{n} && \text{(by (D-9))} \\
&= \sum_{n=1}^{p} \sum_{j=1}^{p} A^{(p)}_{j} \hat{R}^{(p)}_{m+1-n-j} B^{(p)H}_{n} && \text{(by (D-3))} \\
&= \sum_{j=1}^{p} A^{(p)}_{j} \sum_{n=1}^{p} \check{R}^{(p)}_{n-(m+1-j)} B^{(p)H}_{n} && \text{(by (D-9))} \\
&= \sum_{j=1}^{p} A^{(p)}_{j} \hat{R}^{(p)}_{m+1-j} && \text{(by (D-10))} \\
&= \hat{R}^{(p)}_{m+1} && \text{(by (D-3))}.
\end{aligned} \tag{D-11}$$

Therefore we have extended (D-9) by one step, and the proof follows for all $k \ge p+1$ by induction.

Appendix E

RELATIONSHIP OF DETERMINANTS

The forward correlation recursion is given in (52) as

$$\hat{R}_m^{(p)} = \sum_{n=1}^{p} A_n^{(p)} \hat{R}_{m-n}^{(p)} \quad , \; p+1 \le m. \tag{E-1}$$

The z-transform of this sequence is

$$\hat{R}(z) \equiv \sum_{m=p+1}^{\infty} z^{-m} \hat{R}_m^{(p)} = \sum_{n=1}^{p} z^{-n} A_n^{(p)} \sum_{m=p+1}^{\infty} z^{-(m-n)} \hat{R}_{m-n}^{(p)} . \tag{E-2}$$

The inner sum on m can be expressed as (see (53))

$$\sum_{m=p+1}^{p+n} z^{-(m-n)} R_{m-n} + \sum_{m=p+1+n}^{\infty} z^{-(m-n)} \hat{R}_{m-n}^{(p)} \equiv R_n(z) + \hat{R}(z). \tag{E-3}$$

Therefore,

$$\hat{R}(z) = \sum_{n=1}^{p} z^{-n} A_n^{(p)} R_n(z) + \sum_{n=1}^{p} z^{-n} A_n^{(p)} \hat{R}(z) \tag{E-4}$$

or

$$\hat{R}(z) = \left(I - \sum_{n=1}^{p} z^{-n} A_n^{(p)} \right)^{-1} \sum_{n=1}^{p} z^{-n} A_n^{(p)} R_n(z) . \tag{E-5}$$

At the same time, we define the z-transform of the backward correlation recursion as

$$\check{R}(z) \equiv \sum_{m=p+1}^{\infty} z^{-m} \check{R}_{-m}^{(p)} \tag{E-6}$$

and note that, via (62),

$$\check{R}^H(z) = \sum_{m=p+1}^{\infty} z^{-m} \check{R}_{-m}^{(p)H} = \sum_{m=p+1}^{\infty} z^{-m} \hat{R}_m^{(p)} = \hat{R}(z) . \tag{E-7}$$

A comment on notation is timely here. If matrix

$$G(z) \equiv \sum_n z^{-n} D_n \ , \tag{E-8}$$

where z is a complex scalar variable, then

$$G^H(z) \equiv \sum_n z^{-n} D_n^H \ . \tag{E-9}$$

But

$$G(z)^H = \sum_n (z^*)^{-n} D_n^H \ , \tag{E-10}$$

which is not always equal to (E-9), unless z is real.

But let us also develop definition (E-7) by means of backward recursion (55), in a manner similar to that above in (E-1) through (E-5). We find

$$\overset{\vee}{R}^H(z) = \sum_{m=p+1}^{\infty} z^{-m} \overset{\vee}{R}^{(p)H}_{-m} = \sum_{m=p+1}^{\infty} z^{-m} \sum_{n=1}^{p} \overset{\vee}{R}^{(p)H}_{n-m} B_n^{(p)H} \quad (\text{by (55)})$$

$$= \sum_{n=1}^{p} z^{-n} \left(\sum_{m=p+1}^{\infty} z^{-(m-n)} \overset{\vee}{R}^{(p)H}_{n-m} \right) B_n^{(p)H} \ . \tag{E-11}$$

The inner sum on m is

$$\sum_{m=p+1}^{p+n} z^{-(m-n)} R^H_{n-m} + \sum_{m=p+1+n}^{\infty} z^{-(m-n)} \overset{\vee}{R}^{(p)H}_{n-m} = R_n(z) + \overset{\vee}{R}^H(z) \ , \tag{E-12}$$

where we used (56), (2), (E-3), and (E-11). Therefore

$$\overset{\vee}{R}^H(z) = \sum_{n=1}^{p} z^{-n} R_n(z) B_n^{(p)H} + \overset{\vee}{R}^H(z) \sum_{n=1}^{p} z^{-n} B_n^{(p)H} \tag{E-13}$$

or

$$\overset{\vee}{R}^H(z) = \sum_{n=1}^{p} z^{-n} R_n(z) B_n^{(p)H} \left(I - \sum_{n=1}^{p} z^{-n} B_n^{(p)H} \right)^{-1} \tag{E-14}$$

Combining (E-5) and (E-14) according to (E-7), we see that

$$\det\left(I-\sum_{n=1}^{p} z^{-n} A_n^{(p)}\right) \quad \text{and} \quad \det\left(I-\sum_{n=1}^{p} z^{-n} B_n^{(p)H}\right) \tag{E-15}$$

must have the same zeros, since these two quantities determine the singularity locations of (E-5) and (E-14). The quantity $\mathcal{R}_n(z)$ defined in (E-3) is singular only at $z = 0$.

Furthermore

$$\det\left(I-\sum_{n=1}^{p} z^{-n} A_n^{(p)}\right) = z^{-Mp}\det\left(z^p I - z^{p-1}A_1^{(p)} - \cdots - A_p^{(p)}\right) = \frac{\prod_{j=1}^{Mp}(z-z_j)}{z^{Mp}} \tag{E-16}$$

and

$$\det\left(I-\sum_{n=1}^{p} z^{-n} B_n^{(p)H}\right) = z^{-Mp}\det\left(z^p I - z^{p-1}B_1^{(p)H} - \cdots - B_p^{(p)H}\right) = \frac{\prod_{j=1}^{Mp}(z-z_j)}{z^{Mp}}\,, \tag{E-17}$$

where we have utilized the observations that the quantities in (E-15) have the same zeros, the same pole at $z=0$, and the same scale factor. Therefore the two determinants in (E-15) are equal.

Also since

$$\det\left(I-\sum_{n=1}^{p} z^{-n} G_n\right) = 1 - z^{-1}\operatorname{tr} G_1 - \cdots + (-1)^M z^{-Mp}\det G_p\,, \tag{E-18}$$

it follows that

$$\operatorname{tr} A_1^{(p)} = \operatorname{tr} B_1^{(p)H} = \left(\operatorname{tr} B_1^{(p)}\right)^* \tag{E-19}$$

and

$$\det A_p^{(p)} = \det B_p^{(p)H} = \left(\det B_p^{(p)}\right)^*. \tag{E-20}$$

Numerical examples show that generally

$$\left|\operatorname{tr} A_k^{(p)}\right| \neq \left|\operatorname{tr} B_k^{(p)}\right| \quad \text{for} \quad 1 < k \leq p \tag{E-21}$$

and

$$\left|\det A_k^{(p)}\right| \neq \left|\det B_k^{(p)}\right| \quad \text{for} \quad k < p. \tag{E-22}$$

Appendix F

SPECTRUM FROM EXTRAPOLATED CORRELATIONS

The forward correlation recursion is given by

$$\hat{R}_m^{(p)} = \sum_{n=1}^{p} A_n^{(p)} \hat{R}_{m-n}^{(p)} \quad , \quad m \geq p+1 \; , \tag{F-1}$$

where

$$\hat{R}_m^{(p)} = R_m \quad , \quad |m| \leq p \tag{F-2}$$

and

$$\hat{R}_{-m}^{(p)} = \hat{R}_m^{(p)^H} \quad , \; m \geq p+1 . \tag{F-3}$$

We wish to evaluate the z-transform of $\{\Delta \hat{R}_m^{(p)}\}$:

$$\mathcal{G}^{(p)}(z) = \Delta \sum_{m=-\infty}^{\infty} z^{-m} \hat{R}_m^{(p)} \tag{F-4}$$

In order to do so, consider a fictitious process $\{\hat{X}_n\}$ with the correlation given by (F-1) through (F-3) . Consider the output of the optimum predictive error filter, given by

$$\hat{Y}_k \equiv - \sum_{n=0}^{p} A_n^{(p)} \hat{X}_{k-n} \quad , \quad \text{all} \quad k \tag{F-5}$$

The crosscorrelation

$$\hat{C}_m = \overline{\hat{Y}_k \hat{X}_{k-m}^H} = - \sum_{n=0}^{p} A_n^{(p)} \overline{\hat{X}_{k-n} \hat{X}_{k-m}^H} = - \sum_{n=0}^{p} A_n^{(p)} \hat{R}_{n-m}^{(p)} \quad , \text{all } m \tag{F-6}$$

using (F-5) and (F-1), we see that

$$\hat{C}_m = 0 \quad \text{for} \quad m \geq 1 \qquad \left(\hat{C}_m \neq 0 \text{ for } m \leq 0\right) , \tag{F-7}$$

that is, predictive error filter output $\hat{Y}_k$ is uncorrelated with all past values of input $\hat{X}_k$.

Also, output autocorrelation

$$\hat{D}_m = \overline{\hat{Y}_k \hat{Y}_{k-m}^H} = -\sum_{n=0}^{p} \overline{\hat{Y}_k \hat{X}_{k-n}^H} A_n^{(p)H} = -\sum_{n=0}^{p} \hat{C}_{m-n} A_n^{(p)H}, \qquad \text{(F-8)}$$

using (F-5) and (F-6). But now employment of (F-7) in (F-8) shows that

$$\hat{D}_m = 0 \quad \text{for} \quad m \geq 1. \qquad \text{(F-9)}$$

Also (F-7), (F-6), (F-2), and (12) yield

$$\hat{D}_0 = \hat{C}_0 = -\sum_{n=0}^{p} A_n^{(p)} \hat{R}_{-n}^{(p)} = -\sum_{n=0}^{p} A_n^{(p)} R_{-n} = U_p. \qquad \text{(F-10)}$$

And since, from definition (F-8),

$$\hat{D}_{-m} = \hat{D}_m^H, \qquad \text{(F-11)}$$

we have

$$\hat{D}_m = \begin{cases} U_p, & m = 0 \\ 0, & \text{otherwise} \end{cases}; \qquad \text{(F-12)}$$

that is, predictive error filter output $\hat{Y}_k$ is white for input $\hat{X}_k$. (Of course, U_p is not diagonal).

At the same time, autocorrelation $\hat{D}_m$ can be expressed (by means of (F-5)) as

$$\hat{D}_m = \sum_{n=0}^{p} \sum_{j=0}^{p} A_n^{(p)} \overline{\hat{X}_{-n} \hat{X}_{m-j}^H} A_j^{(p)H} = \sum_{n=0}^{p} \sum_{j=0}^{p} A_n^{(p)} \hat{R}_{m-j-n}^{(p)} A_j^{(p)H}, \text{ all } m \qquad \text{(F-13)}$$

Therefore the z-transform of $\{\Delta \hat{D}_m\}$ is

$$\Delta \sum_{m=-\infty}^{\infty} z^{-m} \hat{D}_m = \sum_{n=0}^{p} z^{-n} A_n^{(p)} \Delta \sum_{m=-\infty}^{\infty} z^{-(m+j-n)} \hat{R}_{m+j-n}^{(p)} \sum_{j=0}^{p} z^{j} A_j^{(p)H}$$

$$= \mathcal{A}_A^{(p)}(z)\, \mathcal{G}^{(p)}(z)\, \mathcal{A}_A^{(p)H}(z^{-1}), \tag{F-14}$$

where we have used (F-13), (23), (F-4), and (E-9). When we couple (F-14) with (F-12), we obtain

$$\Delta U_p = \mathcal{A}_A^{(p)}(z)\, \mathcal{G}^{(p)}(z)\, \mathcal{A}_A^{(p)H}(z^{-1}) \tag{F-15}$$

or

$$\mathcal{G}^{(p)}(z) = \Delta \left[\mathcal{A}_A^{(p)}(z)\right]^{-1} U_p \left[\mathcal{A}_A^{(p)H}(z^{-1})\right]^{-1}, \tag{F-16}$$

where matrix U_p is independent of z. This is one of the main results of this appendix.

If we let (for f real)

$$z = \exp(i2\pi f\Delta), \quad |f| < \frac{1}{2\Delta}, \tag{F-17}$$

and denote the forward predictive error filter transfer function and spectral estimate as

$$\mathcal{A}_A^{(p)}(\exp(i2\pi f\Delta)) = -\sum_{n=0}^{p} \exp(-i2\pi f n\Delta) A_n^{(p)} = H_A^{(p)}(f),$$
$$\mathcal{G}^{(p)}(\exp(i2\pi f\Delta)) = \Delta \sum_{m=-\infty}^{\infty} \exp(-i2\pi f m\Delta) \hat{R}_m^{(p)} = G^{(p)}(f), \tag{F-18}$$

respectively, then the spectrum of process $\{\hat{X}_n\}$ can be expressed as

$$G^{(p)}(f) = \Delta\, H_A^{(p)}(f)^{-1}\, U_p\, H_A^{(p)}(f)^{-1^H}, \tag{F-19}$$

where we have utilized the result that (see (E-8)) through (E-10))

$$\left[\mathcal{A}_A^{(p)H}(\exp(-i2\pi f\Delta))\right]^{-1} = \left[\mathcal{A}_A^{(p)}(\exp(i2\pi f\Delta))^H\right]^{-1} = H_A^{(p)}(f)^{H^{-1}} = H_A^{(p)}(f)^{-1^H} \tag{F-20}$$

The procedure for the backward correlation recursion parallels that above to yield (using (23))

$$\mathcal{G}^{(p)}(z) = \Delta \left[\mathcal{A}_B^{(p)}(z)\right]^{-1} V_p \left[\mathcal{A}_B^{(p)H}(z^{-1})\right]^{-1} \tag{F-21}$$

and

$$G^{(p)}(f) = \Delta\ H_B^{(p)}(f)^{-1} V_p\ H_B^{(p)}(f)^{-1H} . \tag{F-22}$$

ZERO LOCATIONS OF $\mathcal{A}_A^{(p)}(z)^{-1}$

Assume that $\mathcal{A}_A^{(p)}(z)^{-1} \equiv Q(z)$ has a zero at $z=z_1 \neq 0$; that is,

$$\text{assume}\quad Q(z_1) = 0,\ \text{the zero matrix}, \tag{F-23}$$

where $0 < |z_1|$. But

$$\mathcal{A}_A^{(p)}(z) = -\sum_{n=0}^{p} z^{-n} A_n^{(p)} = -z^{-p}\left[A_p^{(p)} + z A_{p-1}^{(p)} + \cdots + z^{p-1} A_1^{(p)} - z^p I\right] \tag{F-24}$$

Therefore $\mathcal{A}_A^{(p)}(z_1)$ is finite for $0<|z_1|$, yielding

$$Q(z_1)\,\mathcal{A}_A^{(p)}(z_1) = 0 \neq I. \tag{F-25}$$

Therefore assumption (F-23) is invalid, indicating that

$$Q(z) \neq 0 \quad \text{for} \quad 0 < |z| . \tag{F-26}$$

Now from (F-24)

$$\mathcal{A}_A^{(p)}(z) \sim -z^{-p} A_p^{(p)} \quad \text{as} \quad |z| \to 0 ; \tag{F-27}$$

therefore,

$$Q(z) \sim -z^p A_p^{(p)^{-1}} \quad \text{as} \quad |z| \to 0. \tag{F-28}$$

Thus Q(z) has a p-th order zero at z=0, but is not equal to the zero matrix for $0 < |z|$. Of course, the individual elements of matrix Q(z) can have zeros anywhere.

POLE LOCATIONS OF $\mathcal{A}_A^{(p)}(z)^{-1}$

Since from (F-24)

$$\mathcal{A}_A^{(p)}(z) = -z^{-p} Q_p(z), \tag{F-29}$$

where $Q_p(z)$ is a matrix of polynomials in z of order p, it follows that

$$Q(z) = -z^p Q_p(z)^{-1} = -\frac{z^p}{\det Q_p(z)} \hat{Q}_{(M-1)p}(z), \tag{F-30}$$

where $\hat{Q}_{(M-1)p}(z)$ is a matrix of polynomials in z of order (M-1)p. Therefore the poles of Q(z) are caused by the zeros of det $Q_p(z)$; that is, the poles of $\mathcal{A}_A^{(p)}(z)^{-1}$ are caused by the zeros of det $\mathcal{A}_A^{(p)}(z)$. As $|z| \to \infty$, $\mathcal{A}_A^{(p)}(z) \sim I$ from (F-24); therefore, $Q(z) \sim I$ as $|z| \to \infty$, so that Q(z) has no poles at $|z| = \infty$. Thus the poles of Q(z) are located where det $\mathcal{A}_A^{(p)}(z) = 0$.

We now consider the problem of determining when det $\mathcal{A}_A^{(p)}(z) = 0$: the following derivation is based upon Ref. 7. Let

$$\underline{f}_k = \begin{bmatrix} X_k \\ \\ X_{k-p+1} \end{bmatrix} \tag{F-31}$$

be an Mpxl matrix. Define prediction

$$\hat{\underline{f}}_k = C \underline{f}_{k-1}, \tag{F-32}$$

where C is MpxMp. And define error

$$\delta_k = \mathcal{P}_k - \hat{\mathcal{P}}_k = \mathcal{P}_k - C\mathcal{P}_{k-1} . \tag{F-33}$$

Then

$$\begin{aligned}\overline{\delta_k \delta_k^H} &= \overline{(\mathcal{P}_k - C\mathcal{P}_{k-1})(\mathcal{P}_k^H - \mathcal{P}_{k-1}^H C^H)} \\ &= \mathcal{U}_0 - C\mathcal{U}_{-1} - \mathcal{U}_1 C^H + C\mathcal{U}_0 C^H \\ &= \mathcal{U}_0 - \mathcal{U}_1\mathcal{U}_0^{-1}\mathcal{U}_1^H + (C - \mathcal{U}_1\mathcal{U}_0^{-1})\mathcal{U}_0(C - \mathcal{U}_1\mathcal{U}_0^{-1})^H\end{aligned} \tag{F-34}$$

where

$$\mathcal{U}_m \equiv \overline{\mathcal{P}_k \mathcal{P}_{k-m}^H} . \tag{F-35}$$

The minimum value of $\mathrm{tr}\ \overline{\delta_k \delta_k^H}$ is realized when (see appendix B) we select

$$C = \mathcal{U}_1\mathcal{U}_0^{-1} . \tag{F-36}$$

The corresponding value of

$$\overline{\delta_k \delta_k^H} = \mathcal{U}_0 - \mathcal{U}_1\mathcal{U}_0^{-1}\mathcal{U}_1^H = \mathcal{U}_0 - C\mathcal{U}_0 C^H , \tag{F-37}$$

since $\mathcal{U}_0^H = \mathcal{U}_0$. Now let the left eigenvectors and eigenvalues of the optimum C be denoted as

$$\xi_m^H C = \lambda_m \xi_m^H , \quad 1 \le m \le Mp . \tag{F-38}$$

(The eigenvectors $\{\xi_m\}$ may not all be linearly independent). Then

$$\begin{aligned}0 \le \overline{|\xi_m^H \delta_k|^2} &= \xi_m^H \overline{\delta_k \delta_k^H} \xi_m = \xi_m^H (\mathcal{U}_0 - C\mathcal{U}_0 C^H)\xi_m \\ &= \xi_m^H \mathcal{U}_0 \xi_m (1 - |\lambda_m|^2) .\end{aligned} \tag{F-39}$$

Now $\mathcal{U}_0$ is Hermitian, block Toeplitz, non-negative definite, and has the form

$$\mathcal{U}_0 = \overline{\mathcal{S}_K \mathcal{S}_K^H} = \begin{bmatrix} R_0 & R_1 & \cdots & R_{p-1} \\ R_{-1} & \ddots & & \\ \vdots & & & \\ R_{-p} & & & R_0 \end{bmatrix}. \tag{F-40}$$

Therefore $|\lambda_m| \leq 1$ for $1 \leq m \leq Mp$; that is, all the eigenvalues of C are bounded by unity in magnitude. Furthermore, Ref. 7, p. 134, shows that if there is no exact linear relation between the elements of $X_k, X_{k-1}, \ldots, X_{k-p}$, then $|\lambda_m| < 1$ for $1 \leq m \leq Mp$ (see also appendix B).

Now we develop the error in (F-33) in more detail:

$$\mathcal{S}_K = \mathcal{S}_k - C\mathcal{S}_{k-1} = \begin{bmatrix} X_k \\ \vdots \\ X_{k-p+1} \end{bmatrix} - \begin{bmatrix} C_1 & \cdots & C_p \\ \vdots & & \\ C_{p1} & \cdots & C_{pp} \end{bmatrix} \begin{bmatrix} X_{k-1} \\ \vdots \\ X_{k-p} \end{bmatrix}$$

$$= \begin{bmatrix} X_k - \sum_{m=1}^{p} C_{1m} X_{k-m} \\ \vdots \\ X_{k-p+1} - \sum_{m=1}^{p} C_{pm} X_{k-m} \end{bmatrix}. \tag{F-41}$$

Minimizing $\operatorname{tr} \overline{\mathcal{S}_k \mathcal{S}_k^H}$ can be seen to make C of the form

$$C = \begin{bmatrix} A_1^{(p)} & A_2^{(p)} & \cdots & A_p^{(p)} \\ I & 0 & & 0 \\ 0 & I & 0 & \\ \vdots & & \ddots & \\ 0 & & I & 0 \end{bmatrix} \tag{F-42}$$

Therefore (Ref. 7, eqs. (35) and (36)),

$$\det(C - \lambda I) = (-\lambda)^{Mp} \det \mathcal{A}_M^{(p)}(\lambda). \tag{F-43}$$

If we were to assume that $\det \mathcal{N}_A^{(p)}(z_1)=0$, where $|z_1| \geq 1$, we would have $\det(C-z_1 I)=0$. But this contradicts $|\lambda_m|<1$ for $1 \leq m \leq M_p$. Therefore, the zeros of $\det \mathcal{N}_A^{(p)}(z)$ all lie inside the unit circle; that is, the poles of Q(z) all lie inside the unit circle.

Appendix G

HERMITIAN PROPERTY OF ONE-STEP EXTRAPOLATED CORRELATION MATRIX ESTIMATES

From (78), at the (p-1)th stage, we know that

$$R_{-m}^{H} = \sum_{n=1}^{p-1} R_{n-m}^{H} B_n^{(p-1)H} = \sum_{n=1}^{p-1} R_{m-n} B_n^{(p-1)H} = \sum_{n=1}^{p-1} A_n^{(p-1)} R_{m-n} = R_m \,, \; 1 \le m \le p-1. \qquad \text{(G-1)}$$

Now we start with (94) and express

$$\begin{aligned} R_{-p}^{(p-1)H} &= \sum_{n=1}^{p-1} R_{n-p}^{H} B_n^{(p-1)H} = \sum_{n=1}^{p-1} R_{p-n} B_n^{(p-1)H} \\ &= \sum_{n=1}^{p-1} \sum_{j=1}^{p-1} A_j^{(p-1)} R_{p-n-j} B_n^{(p-1)H} \qquad \text{(by (G-1))} \\ &= \sum_{j=1}^{p-1} A_j^{(p-1)} \sum_{n=1}^{p-1} R_{p-n-j} B_n^{(p-1)H} \\ &= \sum_{j=1}^{p-1} A_j^{(p-1)} R_{p-j} \qquad \text{(by (G-1))} \\ &= R_p^{(p-1)} \qquad \text{(by (98))}. \end{aligned} \qquad \text{(G-2)}$$

Thus, the one-step extrapolated correlation matrix estimates, based on order p-1, are Hermitians of each other.

Appendix H

INTERRELATIONSHIPS OF U_p AND V_p FOR UNKNOWN CORRELATION CASE

We develop the definition (95) as follows:

$$
\begin{aligned}
U_p &= -\sum_{n=0}^{p} A_n^{(p)} R_{-n} = R_0 - \sum_{n=1}^{p-1} A_n^{(p)} R_{-n} - A_p^{(p)} R_{-p} \\
&= R_0 - \sum_{n=1}^{p-1} \left(A_n^{(p-1)} - A_p^{(p)} B_{p-n}^{(p-1)} \right) R_{-n} - A_p^{(p)} \sum_{n=1}^{p} B_n^{(p)} R_{n-p} \quad \text{(by (79A) and (80B))} \\
&= -\sum_{n=0}^{p-1} A_n^{(p-1)} R_{-n} + A_p^{(p)} \sum_{n=1}^{p-1} B_{p-n}^{(p-1)} R_{-n} \\
&\quad - A_p^{(p)} \left[\sum_{n=1}^{p-1} \left(B_n^{(p-1)} - B_p^{(p)} A_{p-n}^{(p-1)} \right) R_{n-p} + B_p^{(p)} R_0 \right] \quad \text{(by (79B))}
\end{aligned}
\tag{H-1}
$$

Now

$$
\sum_{n=1}^{p-1} B_n^{(p-1)} R_{n-p} = \sum_{j=1}^{p-1} B_{p-j}^{(p-1)} R_{-j} \, . \tag{H-2}
$$

Therefore

$$
\begin{aligned}
U_p &= U_{p-1} + A_p^{(p)} B_p^{(p)} \left[\sum_{n=1}^{p-1} A_{p-n}^{(p-1)} R_{n-p} - R_0 \right] \quad \text{(by (95))} \\
&= U_{p-1} + A_p^{(p)} B_p^{(p)} \sum_{n=1}^{p} A_{p-n}^{(p-1)} R_{n-p} \\
&= U_{p-1} - \Delta_p^{(p)} \nabla_p^{(p)} U_{p-1} \quad \text{(by (95))} \\
&= \left(I - A_p^{(p)} B_p^{(p)} \right) U_{p-1} \, .
\end{aligned}
\tag{H-3}
$$

In a similar manner, we can show that

$$
V_p = \left(I - B_p^{(p)} A_p^{(p)} \right) V_{p-1} \tag{H-4}
$$

In order to show that U_p is Hermitian, we recall the constraint (98) and express

$$U_p = U_{p-1} - A_p^{(p)} B_p^{(p)} U_{p-1} = U_{p-1} - A_p^{(p)} V_{p-1}^H A_p^{(p)H} . \tag{H-5}$$

Therefore if $U_{p-1}^H = U_{p-1}$ and $V_{p-1}^H = V_{p-1}$, it immediately follows that

$$U_p^H = U_p . \tag{H-6}$$

Similarly since

$$V_p = V_{p-1} - B_p^{(p)} A_p^{(p)} V_{p-1} = V_{p-1} - B_p^{(p)} U_{p-1}^H B_p^{(p)H} ,$$

it also immediately follows that

$$V_p^H = V_p . \tag{H-7}$$

But properties (H-6) and (H-7) are obviously true for p = 0, because

$$U_0 = V_0 = R_0 = R_0^H . \tag{H-8}$$

Therefore (H-6) and (H-7) are true for all p, by induction.

In order to relate the determinants of U_p and V_p, we express (H-3) and (H-4) as

$$U_p = A_p^{(p)} (A_p^{(p)-1} - B_p^{(p)}) U_{p-1} , \quad V_p = (A_p^{(p)-1} - B_p^{(p)}) A_p^{(p)} V_{p-1} . \tag{H-9}$$

Therefore if $\det U_{p-1} = \det V_{p-1}$, then

$$\det U_p = \det V_p . \tag{H-10}$$

But (H-10) is obviously true for p=0 by (H-8). Therefore (H-10) is true for all p, by induction.

Properties (H-6), (H-7), and (H-10) applied to (98) immediately show that

$$\det B_p^{(p)} = (\det A_p^{(p)})^* . \tag{H-11}$$

Appendix I

MINIMIZATION OF TRACE OF WEIGHTED ERROR MATRICES

We wish to minimize the trace of (123) by choice of matrix G_p. We use the fact that, for square matrices P and Q,

$$tr(PQ) = \sum_{m,n} P_{mn} Q_{nm} = \sum_{n,m} Q_{nm} P_{mn} = tr(QP), \tag{I-1}$$

to express

$$tr\left(\Lambda_{p-1} E_p + \Gamma_{p-1} F_p\right) = tr\Big[\Lambda_{p-1} S_{p-1}^{(yy)} + \Gamma_{p-1} S_{p-1}^{(xx)} - G_p V_{p-1}^{-1} S_{p-1}^{(yx)H} \Lambda_{p-1} - \Gamma_{p-1} S_{p-1}^{(yx)H} U_{p-1}^{-1} G_p$$
$$+ \left(\Lambda_{p-1} G_p V_{p-1}^{-1} S_{p-1}^{(xx)} V_{p-1}^{-1} + U_{p-1}^{-1} S_{p-1}^{(yy)} U_{p-1}^{-1} G_p \Gamma_{p-1} - \Lambda_{p-1} S_{p-1}^{(yx)} V_{p-1}^{-1} - U_{p-1}^{-1} S_{p-1}^{(yx)} \Gamma_{p-1}\right) G_p^H\Big] \tag{I-2}$$

Now (I-2) is an analytic function of the variables $Re(G_{mn})$ and $Im(G_{mn})$. Therefore the minimum of (I-2) is realized simply by setting the coefficient of G_p^H equal to zero (Ref. 20). We obtain, after premultiplying by Λ_{p-1}^{-1} and post-multiplying by Γ_{p-1}^{-1}, the equation for G_p:

$$G_p V_{p-1}^{-1} S_{p-1}^{(xx)} V_{p-1}^{-1} \Gamma_{p-1}^{-1} + \Lambda_{p-1}^{-1} U_{p-1}^{-1} S_{p-1}^{(yy)} U_{p-1}^{-1} G_p = S_{p-1}^{(yx)} V_{p-1}^{-1} \Gamma_{p-1}^{-1} + \Lambda_{p-1}^{-1} U_{p-1}^{-1} S_{p-1}^{(yx)} \tag{I-3}$$

(G_p is not Hermitian or Toeplitz, as numerical examples will show.) In terms of $A_p^{(p)}$ and $B_p^{(p)}$, we have the simultaneous equations

$$A_p^{(p)} S_{p-1}^{(xx)} V_{p-1}^{-1} \Gamma_{p-1}^{-1} + \Lambda_{p-1}^{-1} U_{p-1}^{-1} S_{p-1}^{(yy)} B_p^{(p)H} = S_{p-1}^{(yx)} V_{p-1}^{-1} \Gamma_{p-1}^{-1} + \Lambda_{p-1}^{-1} U_{p-1}^{-1} S_{p-1}^{(yx)} \tag{I-4}$$

$$A_p^{(p)} V_{p-1} - U_{p-1} B_p^{(p)H} = 0$$

where we utilized (122).

Appendix J

COMPUTATION OF FILTER TRANSFER FUNCTION

The forward predictive error filter transfer function is given in (68) as

$$H_A^{(p)}(f) = -\sum_{n=0}^{p} \exp(-i\,2\pi f n\Delta)\, A_n^{(p)} \quad , \; |f| < \frac{1}{2\Delta} \tag{J-1}$$

Now divide the frequency range $(-\frac{1}{2\Delta}, \frac{1}{2\Delta})$ into N_F cells of width

$$\Delta_f = \frac{1}{N_F}\,\frac{1}{\Delta} \; . \tag{J-2}$$

Then for $|m| \le N_F/2$,

$$\begin{aligned} H_A^{(p)}(m\,\Delta_f) = H_A^{(p)}\left(\frac{m}{N_F\Delta}\right) &= -\sum_{n=0}^{p} \exp(-i\,2\pi mn/N_F)\, A_n^{(p)} \\ &= \sum_{n=0}^{N_F-1} \exp(-i\,2\pi mn/N_F)\, Z_n \; , \end{aligned} \tag{J-3}$$

where

$$Z_n = \begin{cases} -A_n^{(p)} \; , & 0 \le n \le p \\ 0, & p+1 \le n \le N_F-1 \end{cases} \tag{J-4}$$

Now if we let the sum in (J-3) be denoted as an N_F-point FFT,

$$\sum_{n=0}^{N_F-1} \exp(-i\,2\pi mn/N_F) Z_n = \tilde{Z}_m \; , \quad 0 \le m \le N_F-1, \tag{J-5}$$

then (J-3) becomes

$$H_A^{(p)}\left(\frac{m}{N_F\Delta}\right) = \begin{cases} \tilde{Z}_m \; , & 0 \le m \le N_F/2 \\ \tilde{Z}_{N_F+m} \; , & -N_F/2 \le m \le -1 \end{cases} \; . \tag{J-6}$$

Then quantity $\tilde{Z}_m$ in (J-5) is an MxM matrix for each value of m.

Appendix K

PROGRAM FOR SPECTRAL ANALYSIS

In this appendix we present the program for the procedure summarized in (184) - (187) and (165). The spectral estimate, (165), is computed at frequencies $\{m/(N_F\Delta)\}$.

$$G^{(p)}\left(\frac{m}{N_F\Delta}\right) = \Delta\, H_A^{(p)}\left(\frac{m}{N_F\Delta}\right)^{-1} U_p \left[H_A^{(p)}\left(\frac{m}{N_F\Delta}\right)^{-1}\right]^H , \quad |m| \le N_F/2, \tag{K-1}$$

where the forward predictive error transfer function $H_A^{(p)}\left(\frac{m}{N_F\Delta}\right)$ is given by (J-6). The specific scaling adopted is based upon (166), which takes the sampled form

$$\sum_{m=-N_F/2}^{N_F/2} w_m \Delta_f G^{(p)}(m\Delta_f) \cong R_0 , \tag{K-2}$$

where $\{w_m\}$ is a set of integration weights (e.g., trapezoidal). The approximation is a good one if $G^{(p)}(f)$ is sampled finely enough; that is, if N_F is large enough to resolve the peaks and valleys of $G^{(p)}(f)$. If we employ (J-2), (K-2) becomes

$$\sum_{m=-N_F/2}^{N_F/2} w_m \frac{1}{N_F\Delta} G^{(p)}\left(\frac{m}{N_F\Delta}\right) \cong R_0 ; \tag{K-3A}$$

or, for trapezoidal weighting,

$$\sum_{m=-\frac{N_F}{2}+1}^{N_F/2} \frac{1}{N_F\Delta} G^{(p)}\left(\frac{m}{N_F\Delta}\right) \cong R_0 , \tag{K-3B}$$

where we have employed the periodic nature of $G^{(p)}(f)$ (See (165) and (68)).

Thus the sum of samples $\frac{1}{N_F\Delta} G^{(p)}\left(\frac{m}{N_F\Delta}\right)$ equals the sample power, (80).

For a real multivariate process, we can employ (171); a modified form emerges:

$$\mathrm{Re} \sum_{m=0}^{N_F/2} \tilde{w}_m \frac{2}{N_F\Delta} G^{(p)}\left(\frac{m}{N_F\Delta}\right) \equiv R_0 \quad \text{for real process,} \tag{K-4}$$

where $\{\tilde{w}_m\}$ is another set of integration weights. This is the form programmed in the following; the quantities computed are

$$\frac{2}{N_F\Delta} G^{(p)}\left(\frac{m}{N_F\Delta}\right) = \frac{2}{N_F} H_A^{(p)}\left(\frac{m}{N_F\Delta}\right)^{-1} U_p \left[H_A^{(p)}\left(\frac{m}{N_F\Delta}\right)^{-1}\right]^H \quad \text{for } 0 \le m \le \frac{N_F}{2} . \tag{K-5}$$

The real part of their weighted sum equals the sample power, R_0. The FFT used here to evaluate (J-5) is given in Ref. 21; it is limited to powers of 2, but could be replaced if desired. Input parameters are N, PMAX, and NF in line 22, and the input data call is in line 37 and SUBROUTINE DATA; all these quantities have to be changed by the user to fit his particular application. The program is written for a real multivariate process (general M), with the exception of FUNCTION DETERM, SUBROUTINEs SDM, INVERT, and SOLVE, and the printout of the spectral density matrix, (K-5). Arrays used in the program are explained by comment statements. A sample printout follows the program.

```
C  MULTIVARIATE LINEAR PREDICTIVE SPECTRAL ANALYSIS,
C  EMPLOYING WEIGHTED FORWARD AND BACKWARD AVERAGING.
C  THIS PROGRAM IS WRITTEN FOR REAL PROCESSES AND GENERAL M, WITH THE
C  EXCEPTION OF FUNCTION DETERM AND SUBROUTINES SDM, INVERT, AND SOLVE,
C  AND THE PRINT OUT OF THE SPECTRAL DENSITY MATRIX.
C  USER: CHANGE LINES 22 AND 37, AND REPLACE SUBROUTINE DATA.
C  M = DIMENSIONALITY OF MULTIVARIATE PROCESS; INTEGER INPUT
C  N = NUMBER OF DATA POINTS IN EACH PROCESS; INTEGER INPUT
C  X(1,1)...X(N,1),...,X(1,M)...X(N,M) = INPUT DATA; ALTERED ON OUTPUT
C  PMAX = MAXIMUM ORDER OF FILTER; INTEGER INPUT
C  NF = SIZE OF FFT (MUST BE A POWER OF 2 TO USE WKLFFT); INTEGER INPUT
C  AVE = MEANS OF INPUT DATA; OUTPUT
C  R = COVARIANCE MATRIX OF INPUT DATA; OUTPUT
C  AIC = AKAIKE'S INFORMATION CRITERION; OUTPUT
C  PBEST = BEST ORDER OF FILTER; INTEGER OUTPUT
C  UBEST = MATRIX OF COEFFICIENTS IN SPECTRAL ESTIMATE; OUTPUT
C  AP = MATRIX OF FORWARD PARTIAL CORRELATION COEFFICIENTS; THEN =
C  MATRIX OF FORWARD PREDICTIVE FILTER COEFFICIENTS FOR PBEST; OUTPUT
C  BP = MATRIX OF BACKWARD PARTIAL CORRELATION COEFFICIENTS; OUTPUT
C  XX,YY = SPECTRAL MATRICES; OUTPUT
      PARAMETER M=2     ! BIVARIATE PROCESS
      PARAMETER N= 100 , PMAX= 10, NF=1024, NF41=NF/4+1
      INTEGER PBEST,P
      DIMENSION X(N,M),Y(N,M),Z(N,M),UBEST(M,M),AP(M,M,PMAX),
     $BP(M,M,PMAX),AVE(M),XX(NF,M,M),YY(NF,M,M),COSI(NF41),
     $U(M,M),V(M,M),UI(M,M),VI(M,M),A(M,M),B(M,M),R(M,M),
     $WA(M,M),WB(M,M),WC(M,M),WD(M,M),WE(M,M),AIC(PMAX),AICO(2)
      EQUIVALENCE (X,Y),(AIC(1),AICO(2))
C  PRINT OUT VALUES OF PARAMETERS
      I=N
      J=PMAX
      K=M
      L=NF
      PRINT 1, I,J,K,L
1     FORMAT(1H1,' N =',I6,10X,'PMAX =',I4,10X,'M =',I2,10X,'NF =',I5)
C  INPUT DATA IN X(1,1),..X(N,1),...,X(1,M)...X(N,M)
      CALL DATA
      PRINT 2
2     FORMAT(/' INPUT DATA:')
      J=N-99
      L=N-200
      DO 3 I=1,M
      PRINT 4, I
      IF(N.LE.200) GO TO 5
      PRINT 6, (X(K,I),K=1,100)
      PRINT 7, L
7     FORMAT(I6,' INPUT DATA POINTS NOT PRINTED HERE')
      PRINT 6, (X(K,I),K=J,N)
      GO TO 3
5     PRINT 6, (X(K,I),K=1,N)
3     CONTINUE
4     FORMAT('   PROCESS NUMBER',I2)
6     FORMAT(5E20.8)
C  EVALUATE PARTIAL CORRELATION COEFFICIENTS
      CALL PCC
      PRINT 8
```

```
8     FORMAT(/' MEANS OF INPUT DATA:')
      PRINT 6, (AVE(I),I=1,M)
      PRINT 9
9     FORMAT(/' COVARIANCE MATRIX OF INPUT DATA:')
      PRINT 6, ((R(I,J),I=1,M),J=1,M)
      PRINT 10
10    FORMAT(/' AKAIKE INFORMATION CRITERION:')
      PRINT 11, (P,AIC(P),P=0,PMAX)
11    FORMAT(I10,E20.8)
      PRINT 12, PBEST
12    FORMAT(/' PBEST =',I3)
      PRINT 13
13    FORMAT(/' UBEST:')
      PRINT 6, ((UBEST(I,J),I=1,M),J=1,M)
      PRINT 14
14    FORMAT(/' FORWARD PARTIAL CORRELATION COEFFICIENTS:')
      DO 15 P=1,PMAX
15    PRINT 16, P,((AP(I,J,P),I=1,M),J=1,M)
16    FORMAT(I10,6E20.8)
      PRINT 17
17    FORMAT(/' BACKWARD PARTIAL CORRELATION COEFFICIENTS:')
      DO 18 P=1,PMAX
18    PRINT 16, P,((BP(I,J,P),I=1,M),J=1,M)
      IF(PBEST.EQ.0) GO TO 19
C  EVALUATE PREDICTIVE FILTER COEFFICIENTS
      CALL PFC
      PRINT 20
20    FORMAT(/' FORWARD PREDICTIVE FILTER COEFFICIENTS FOR PBEST:')
      DO 21 P=1,PBEST
21    PRINT 16, P,((AP(I,J,P),I=1,M),J=1,M)
C  EVALUATE PREDICTIVE-ERROR FILTER TRANSFER FUNCTION
19    CALL PEFTF
C  EVALUATE SPECTRAL DENSITY MATRIX AND COHERENCE
      K=NF/2+1
      CALL SUM
      PRINT 22
22    FORMAT(/' SPECTRAL DENSITY MATRIX AND COHERENCE FOR M=2:')
      PRINT 23
23    FORMAT(8X,'BIN',10X,'AUTO11',14X,'AUTO22',10X,'REAL(CROSS12)',7X,'
     $IMAG(CROSS12)',9X,'MAG SQ COH',11X,'ARGUMENT')
      PRINT 16, (L,XX(L,1,1),XX(L,2,2),XX(L,1,2),YY(L,1,2),YY(L,1,1),YY(
     $L,2,2), L=1,K)
      SUBROUTINE DATA
C  THIS SUBROUTINE GENERATES DATA FOR M=2, BIVARIATE PROCESS
      DEFINE IRAND=I*5**15+((1-SIGN(1,I*5**15))/2)*34359738367
      DEFINE RAND=FLOAT(I)/34359738367.
      I=5281
      TA=0.
      TB=0.
      DO 1 K=1,100   @ WILL DISCARD THESE INITIAL POINTS
      I=IRAND
      T=.85*TA-.75*TB+RAND-.5
      I=IRAND
```

```
      TB=.65*TA+.55*TB+RAND-.5
1     TA=T
      X(1,1)=TA
      X(1,2)=TB
      DO 2 K=2,N
      I=IRAND
      T=.85*TA-.75*TB+RAND+.5
      I=IRAND
      TB=.65*TA+.55*TB+RAND-.5
      TA=T
      X(K,1)=TA
2     X(K,2)=TB
      RETURN
      SUBROUTINE PCC
C  THIS SUBROUTINE COMPUTES PBEST, UBEST, AND THE PARTIAL
C  CORRELATION COEFFICIENTS FOR P = 1 TO PMAX; ANY M
      I=N
      J=PMAX
      IA=3.*SQRT(N)/M
      IF(PMAX.GT.IA) PRINT 1, J,I,IA
1     FORMAT(/' PMAX =',I4,' IS TOO LARGE FOR NUMBER OF DATA POINTS N ='
     $,I5,';  SEARCH LIMITED TO P =',I4)
      IA=MIN(IA,PMAX)          ;  UPPER BOUND ON PMAX; EQ 183
      FAC=2.*M*M/N             ;  FAC=0.  WOULD FORCE PBEST EQUAL TO PMAX
C  SUBTRACT MEANS; FILL IN DATA ARRAYS; EQ 110
      DO 2 I=1,M
      TA=0.
      DO 3 K=1,N
3     TA=TA+Y(K,I)
      TA=TA/N
      AVE(I)=TA
      DO 2 K=1,N
      Y(K,I)=Y(K,I)-TA
2     Z(K,I)=Y(K,I)
C  INITIALIZE CORRELATION MATRICES; EQS 82, 114, AND 105
      CALL AUTO(2,N-1,Y,WC)
      DO 4 I=1,M
      DO 4 J=I,M
      TA=Y(1,I)*Y(1,J)
      TB=Y(N,I)*Y(N,J)
      R(I,J)=(WC(I,J)+TA+TB)/N
      WA(I,J)=WC(I,J)+TB
      WB(I,J)=WC(I,J)+TA
      R(J,I)=R(I,J)
      WA(J,I)=WA(I,J)
4     WB(J,I)=WB(I,J)
      CALL EQUAL(R,U)
      CALL EQUAL(R,V)
      CALL CROSS(2,N,Y,Y,WC)
C  BEGIN RECURSION
      AIC(0)=LOG(DETERM(U))
      AICMIN=AIC(0)
      PBEST=0
```

```
      CALL EQUAL(U,UBEST)
      DO 5 P=1,IA
C  EVALUATE MATRICES REQUIRED IN BILINEAR MATRIX EQUATION; EQ 126
      CALL INVERT(V,VI)
      CALL MULT(VI,WB,WD)
      CALL EQUAL(WD,WB)
      CALL INVERT(U,UI)
      CALL EQUAL(WA,WD)
      CALL MULT(WD,UI,WA)
      CALL AUD(WC,WC,WC)
C  SOLVE BILINEAR MATRIX EQUATION; EQS 157-161
      CALL SOLVE
C  EVALUATE PARTIAL CORRELATION COEFFICIENTS; EQ 124
      CALL MULT(WC,VI,A)
      CALL TRANS(WC,WD)
      CALL MULT(WD,UI,B)
      CALL EQUAL(A,AP(1,1,P))
      CALL EQUAL(B,BP(1,1,P))
C  UPDATE MATRICES U AND V; EQ 181
      CALL MULT(A,WD,WE)
      CALL SUB(U,WE,U)
      CALL MULT(B,WC,WE)
      CALL SUB(V,WE,V)
C  CALCULATE AKAIKE'S INFORMATION CRITERION; EQ 180
      AIC(P)=LOG(DETERM(U))+FAC*P
      IF(AIC(P).GE.AICMIN) GO TO 6
      AICMIN=AIC(P)
      PBEST=P
      CALL EQUAL(U,UBEST)
6     IF(P.EQ.IA) GO TO 5
C  UPDATE DATA SEQUENCES Y AND Z; EQ 111
      L=P+1
      DO 7 K=N,L,-1
      DO 8 I=1,M
      TA=Z(K-1,I)
      DO 9 J=1,M
9     TA=TA-B(I,J)*Y(K,J)
8     Z(K,I)=TA
      DO 10 I=1,M
      TA=Y(K,I)
      DO 11 J=1,M
11    TA=TA-A(I,J)*Z(K-1,J)
10    Y(K,I)=TA
7     CONTINUE
C  CALCULATE NEW CORRELATION MATRICES; EQ 114
      CALL AUTO(P+2,N,Y,WA)
      CALL AUTO(P+1,N-1,Z,WB)
      CALL CROSS(P+2,N,Y,Z,WC)
5     CONTINUE
      IF(M.EQ.1) RETURN
      K=M-1
      DO 12 I=1,K
      L=I+1
      DO 12 J=L,M
```

```
      UBEST(I,J)=.5*(UBEST(I,J)+UBEST(J,I))
12    UBEST(J,I)=UBEST(I,J)
      RETURN
C
      SUBROUTINE PFC
C  THIS SUBROUTINE COMPUTES THE PREDICTIVE
C  FILTER COEFFICIENTS; ANY M; EQ 79
      IF(PBEST.LE.1) RETURN
      DO 1 P=2,PBEST
      IA=P-1
      DO 2 L=1,IA
      IB=P-L
      CALL MULT(AP(1,1,P),BP(1,1,IB),WA)
      CALL SUB(AP(1,1,L),WA,WA)
      CALL MULT(BP(1,1,P),AP(1,1,L),WB)
      CALL SUB(BP(1,1,IB),WB,BP(1,1,IB))
2     CALL EQUAL(WA,AP(1,1,L))
1     CONTINUE
      RETURN
C
      SUBROUTINE PEFTF
C  THIS SUBROUTINE COMPUTES THE PREDICTIVE-ERROR
C  FILTER TRANSFER FUNCTION; ANY M; EQS 68 AND (J-3)-(J-6)
      K=1.4427*LOG(NF)+.5
      CALL QTRCOS(COSI,NF)
      DO 1 I=1,M
      DO 1 J=1,M
      XX(1,I,J)=0.
      IF(I.EQ.J) XX(1,I,J)=1.
      YY(1,I,J)=0.
      IF(PBEST.EQ.0) GO TO 2
      IA=PBEST+1
      DO 3 L=2,IA
      XX(L,I,J)=-AP(I,J,L-1)
3     YY(L,I,J)=0.
2     IA=PBEST+2
      DO 4 L=IA,NF
      XX(L,I,J)=0.
4     YY(L,I,J)=0.
1     CALL MKLFFT(XX(1,1,J),YY(1,I,J),COSI,K,-1)
      RETURN
C
      SUBROUTINE SDM
C  THIS SUBROUTINE COMPUTES THE SPECTRAL DENSITY
C  MATRIX AND COHERENCE FOR M=2; EQS 178 AND K-5
      T=2./NF
      DO 1 L=1,K
      WA(1,1)=XX(L,2,2)
      WA(1,2)=-XX(L,1,2)
      WA(2,1)=-XX(L,2,1)
      WA(2,2)=XX(L,1,1)
      WB(1,1)=YY(L,2,2)
      WB(1,2)=-YY(L,1,2)
      WB(2,1)=-YY(L,2,1)
      WB(2,2)=YY(L,1,1)
```

```
      TA=DETERM(WA)-DETERM(WB)
      TB=WA(1,1)*WB(2,2)+WA(2,2)*WB(1,1)-WA(1,2)*WB(2,1)-WA(2,1)*WB(1,2)
      TA=T/(TA**2+TB**2)
      CALL TRANS(WA,WC)
      CALL MULT(UBEST,WC,WD)
      CALL MULT(WB,WD,WC)
      TB=WC(1,2)-WC(2,1)
      CALL MULT(WA,WD,WC)
      CALL TRANS(WB,WD)
      CALL MULT(UBEST,WD,WE)
      CALL MULT(WB,WE,WD)
      CALL ADD(WC,WD,WC)
      YY(L,1,1)=(WC(1,2)**2+TB**2)/(WC(1,1)*WC(2,2))     * MAG SQ COH
      YY(L,2,2)=ATAN2(TB,WC(1,2))                        * ARGUMENT
      XX(L,1,1)=TA*WC(1,1)                               * AUTO11
      XX(L,2,2)=TA*WC(2,2)                               * AUTO22
      XX(L,1,2)=TA*WC(1,2)                               * REAL(CROSS1_)
      YY(L,1,2)=TA*TB                                    * IMAG(CROSS1 )
      XX(L,2,1)=0.
      YY(L,2,1)=0.
1     CONTINUE
      RETURN
C
      SUBROUTINE CROSS(N1,N2,A,B,C)     @ A,B,A NG
C  THIS SUBROUTINE COMPUTES A CROSS CORRELATION MATRIX; ANY M EQ 11,F
      DIMENSION A(N,M),B(N,M),C(M,M)
      DOUBLE PRECISION T
      DO 1 I=1,M
      DO 1 J=1,M
      T=0.D0
      DO 2 K=N1,N2
2     T=T+A(K,I)*B(K-1,J)
1     C(I,J)=T
      RETURN
C
      SUBROUTINE AUTO(N1,N2,A,B)    @  A,A NG
C  THIS SUBROUTINE COMPUTES AN AUTO CORRELATION MATRIX; ANY M EQ 11,A
      DIMENSION A(N,M),B(M,M)
      DOUBLE PRECISION T
      DO 1 I=1,M
      DO 1 J=I,M
      T=0.D0
      DO 2 K=N1,N2
2     T=T+A(K,I)*A(K,J)
      B(I,J)=T
1     B(J,I)=B(I,J)
      RETURN
C
      SUBROUTINE EQUAL(A,B)
C  THIS SUBROUTINE SETS TWO MXM MATRICES EQUAL
      DIMENSION A(M,M),B(M,M)
      DO 1 I=1,M
      DO 1 J=1,M
1     B(I,J)=A(I,J)
      RETURN
```

```
      SUBROUTINE TRANS(A,B)      @  A,A NG
C  THIS SUBROUTINE TRANSPOSES AN MXM MATRIX
      DIMENSION A(M,M),B(M,M)
      DO 1 I=1,M
      DO 1 J=1,M
1     B(I,J)=A(J,I)
      RETURN
C
      SUBROUTINE ADD(A,B,C)      @  A,B,A OK
C  THIS SUBROUTINE ADDS TWO MXM MATRICES
      DIMENSION A(M,M),B(M,M),C(M,M)
      DO 1 I=1,M
      DO 1 J=1,M
1     C(I,J)=A(I,J)+B(I,J)
      RETURN
C
      SUBROUTINE SUB(A,B,C)      @  A,B,A OK
C  THIS SUBROUTINE SUBTRACTS TWO MXM MATRICES
      DIMENSION A(M,M),B(M,M),C(M,M)
      DO 1 I=1,M
      DO 1 J=1,M
1     C(I,J)=A(I,J)-B(I,J)
      RETURN
C
      SUBROUTINE MULT(A,B,C)     @  A,B,A NG
C  THIS SUBROUTINE MULTIPLIES TWO MXM MATRICES
      DIMENSION A(M,M),B(M,M),C(M,M)
      DO 1 I=1,M
      DO 1 J=1,M
      T=0.
      DO 2 K=1,M
2     T=T+A(I,K)*B(K,J)
1     C(I,J)=T
      RETURN
C
      SUBROUTINE INVERT(A,B)     @  A,A NG
C  THIS SUBROUTINE INVERTS A 2X2 MATRIX
      DIMENSION A(2,2),B(2,2)
      TA=1./DETERM(A)
      B(1,1)=A(2,2)*TA
      B(2,2)=A(1,1)*TA
      B(1,2)=-A(1,2)*TA
      B(2,1)=-A(2,1)*TA
      RETURN
C
      SUBROUTINE SOLVE
C  THIS SUBROUTINE SOLVES BILINEAR MATRIX EQUATION
C  FOR M=2, BIVARIATE PROCESS; EQS 157, 158, AND 162
      TA=WA(1,1)+WA(2,2)+WB(1,1)+WB(2,2)
      TB=DETERM(WA)-DETERM(WB)
      CALL MULT(WC,WB,WD)
      WE(1,1)=WA(2,2)
      WE(1,2)=-WA(1,2)
      WE(2,1)=-WA(2,1)
      WE(2,2)=WA(1,1)
```

```
      CALL MULT(WE,WC,WA)
      CALL ADD(WD,WA,WD)
      WB(1,1)=TA*WB(1,1)+TB
      WB(2,2)=TA*WB(2,2)+TB
      WB(1,2)=TA*WB(1,2)
      WB(2,1)=TA*WB(2,1)
      CALL INVERT(WB,WE)
      CALL MULT(WD,WE,WC)
      RETURN
C
      FUNCTION DETERM(A)
C  THIS FUNCTION COMPUTES THE DETERMINANT OF A 2X2 MATRIX
      DIMENSION A(2,2)
      DETERM=A(1,1)*A(2,2)-A(1,2)*A(2,1)
      RETURN
      END

  SUBROUTINE MKLFFT(X,Y,CC,M,ISN)
  DIMENSION X(1),Y(1),CC(1),L(12)
  EQUIVALENCE (L12,L(1)),(L11,L(2)),(L10,L(3)),(L9,L(4)),(L8,L(5)),
 1(L7,L(6)),(L6,L(7)),(L5,L(8)),(L4,L(9)),(L3,L(10)),(L2,L(11)),
 2(L1,L(12))
  N=2**M
  ND4=N/4
  ND4P1=ND4+1
  ND4P2=ND4P1+1
  ND2P2=ND4+ND4P2
  DO 8 LO=1,M
  LMX=2**(M-LO)
  LIX=2*LMX
  ISCL=N/LIX
  DO 8 LM=1,LMX
  IARG=(LM-1)*ISCL+1
  IF(IARG.LE.ND4P1) GO TO 4
  C=-CC(ND2P2-IARG)
  S=ISN*CC(IARG-ND4)
  GO TO 6
4 C=CC(IARG)
  S=ISN*CC(ND4P2-IARG)
6 DO 8 LI=LIX,N,LIX
  J1=LI-LIX+LM
  J2=J1+LMX
  T1=X(J1)-X(J2)
  T2=Y(J1)-Y(J2)
  X(J1)=X(J1)+X(J2)
  Y(J1)=Y(J1)+Y(J2)
  X(J2)=C*T1-S*T2
  Y(J2)=C*T2+S*T1
8 CONTINUE
  DO 40 J=1,12
  L(J)=1
  IF(J-M) 31,31,40
31 L(J)=2**(M+1-J)
```

```
40 CONTINUE
   JN=1
   DO 60 J1=1,L1
   DO 60 J2=J1,L2,L1
   DO 60 J3=J2,L3,L2
   DO 60 J4=J3,L4,L3
   DO 60 J5=J4,L5,L4
   DO 60 J6=J5,L6,L5
   DO 60 J7=J6,L7,L6
   DO 60 J8=J7,L8,L7
   DO 60 J9=J8,L9,L8
   DO 60 J10=J9,L10,L9
   DO 60 J11=J10,L11,L10
   DO 60 JR=J11,L12,L11
   IF(JN-JR) 51,51,52
51 R=X(JN)
   X(JN)=X(JR)
   X(JR)=R
   FI=Y(JN)
   Y(JN)=Y(JR)
   Y(JR)=FI
52 JN=JN+1
60 CONTINUE
   RETURN
   END

      SUBROUTINE QTRCOS(C,N)
      DIMENSION C(1)
      N41=N/4+1
      SCL=6.283185307/N
      DO 1 I=1,N41
 1    C(I)=COS((I-1)*SCL)
      RETURN
      END
```

N = 100 PMAX = 10 M = 2 NF = 1024

INPUT DATA:

PROCESS NUMBER 1

.53901729+00	.24572077+00	-.48246256+00	-.99235174+00	-.62436315+00
.31143945+00	.74571225+00	.94372392+00	.15172541+00	-.76859383+00
-.13781174+00	.25836023+00	.79770139+00	.71260673+00	.80128585+00
.43642798+00	-.44703931-01	-.71200128+00	-.11368756+01	-.92819643+00
-.33429637+00	.57812569+00	.13680694+01	.10164690+01	.50713280+00
-.10550485+01	-.20570143+01	-.19485568+01	-.14135770+01	-.19008934+00
.18386345+01	.31438660+01	.21384055+01	-.46009872+00	-.23113293+01
-.36081896+01	-.28417680+01	-.23547348+00	.16677773+01	.33429194+01
.28877570+01	.56888448+00	-.22631766+01	-.33943964+01	-.26797284+01
-.43547171+00	.21718071+01	.35407788+01	.23288262+01	-.22656890+00
-.25376441+01	-.32986798+01	-.27345279+01	-.40607340+00	.25071375+01
.44460779+01	.32127956+01	.91031507+00	-.13406594+01	-.27867314+01
-.25558064+01	-.11135026+01	.95231916+00	.31850186+01	.37061642+01
.17390409+01	-.89170616+00	-.24854088+01	-.27824243+01	-.13415888+01
.62164557+00	.30034332+01	.35482952+01	.26236263+01	-.89912541-01
-.34330869+01	-.47749634+01	-.34963096+01	-.69029214+00	.24394748+01
.40683678+01	.36433851+01	.54673234+00	-.18884842+01	-.32592577+01
-.28174284+01	-.13084519+01	.12478004+01	.28549910+01	.27696476+01
.13411797+01	-.16488339+01	-.35090672+01	-.33495106+01	-.18694842+01
.17976016+00	.23938473+01	.25152508+01	.16395330+01	-.34817300+00

PROCESS NUMBER 2

.84303659+00	.86934820+00	.16071017+00	.14386062+00	-.81120574+00
-.96936302+00	-.68125464-01	.71946396+00	.53986827+00	-.30419506-01
-.85995505+00	-.32430927+00	.10317814+00	.34722114+00	.80690885+00
.57350412+00	.39259374+00	.14101283+00	.84664986-01	-.19283811+00
-.11813866+01	-.86226903+00	-.20097572+00	.66497141+00	.14060706+01
.11701580+01	.81481338-01	-.81963249+00	-.17308412+01	-.21933125+01
-.17339195+01	.58163615+00	.24263104+01	.29041306+01	.17109564+01
-.18242586+00	-.25142254+01	-.28079124+01	-.19512028+01	.50994501+00
.25400578+01	.30085762+01	.18208549+01	-.64529851+00	-.28813978+01
-.31391280+01	-.17331168+01	.92467879+00	.25048586+01	.30233470+01
.19451272+01	-.49972244+00	-.29138998+01	-.33593366+01	-.25352063+01
.43625407+00	.27926130+01	.31353639+01	.25105703+01	.87293217+00
-.12635761+01	-.25078586+01	-.25748112+01	-.99730413+00	.13507358+01
.26830553+01	.23373863+01	.82668857+00	-.11139500+01	-.21637863+01
-.25236639+01	-.10761964+01	.98666745+00	.30014957+01	.38264174+01
.23522349+01	-.81600191+00	-.32755376+01	-.45058461+01	-.25775154+01
-.21895836+00	.29419152+01	.36481336+01	.28266277+01	-.60319692-01
-.19925641+01	-.27615544+01	-.26589665+01	-.64737901+00	.17339140+01
.31629709+01	.26245975+01	.38355845+00	-.19680776+01	-.27634547+01
-.24165064+01	-.10100275+01	.11536021+01	.19672252+01	.16684788+01

MEANS OF INPUT DATA:

.12088474-01	.29368686-01

COVARIANCE MATRIX OF INPUT DATA:

.46218387+01	.91577268+00	.37956951+01

AKAIKE INFORMATION CRITERION:

0	.28156752+01
1	-.47116107+01
2	-.46664973+01
3	-.46229553+01

4	-.46316842+01
5	-.45650773+01
6	-.45422159+01
7	-.45357835+01
8	-.44974548+01
9	-.44907190+01
10	-.44389324+01

PBEST = 1

UBEST:

.89002132-01 -.79014897-03 -.79014897-03 .93252867-01

FORWARD PARTIAL CORRELATION COEFFICIENTS:

1	.87150693+00	.63431677+00	-.77024333+00	.56034775+00
2	-.24060294-01	.15673790+00	-.46089508-01	-.13923085+00
3	.12915299+00	.10394043+00	-.69120812-01	-.12937856+00
4	-.11394597+00	.22871864+00	.18147939+00	.72523244-02
5	.39750100-01	.27629165-01	-.10585008+00	.16504833-01
6	.24732766+00	-.11105661-01	-.13494856-01	-.36488586-02
7	.18216120+00	.10071856+00	.86155114-01	-.10703941+00
8	.19121648+00	.93589546-01	-.25144000-01	.10180847-01
9	.82219786-01	.12730163+00	-.10475686+00	-.22288333+00
10	-.53422038-01	.16704631+00	-.42075946-01	-.26114821-01

BACKWARD PARTIAL CORRELATION COEFFICIENTS:

1	.56612993+00	-.63141833+00	.77098253+00	.66572675+00
2	-.38369942-01	-.66131392-01	.95437832-01	-.11108892+00
3	.97749089-01	-.38619501-01	.81535051-01	-.11780613+00
4	-.33532630-01	.18090538+00	.21567873+00	.98908506-01
5	-.22097769-02	-.11387029+00	.31994226-01	.28316177-01
6	.22834081+00	.77628009-01	-.75839960-02	-.71869090-02
7	.20997133+00	.19255909+00	.53676562-01	-.89760510-01
8	.17425043+00	.38450344-01	.85870568-01	.43625234-01
9	.39914736-01	-.11974495+00	.33887566-01	-.22667238+00
10	-.58663283-01	-.74213017-01	.13284003+00	.24026474-01

FORWARD PREDICTIVE FILTER COEFFICIENTS FOR PBEST:

1	.87150693+00	.63431677+00	-.77024333+00	.56034775+00

SPECTRAL DENSITY MATRIX AND COHERENCE FOR M=2:

BIN	AUTO11	AUTO22	REAL(CROSS12)	IMAG(CROSS12)	MAG SQ COH	ARGUMENT
1	.48031416-03	.24469207-03	.10474293-03	.00000000	.93347806-01	.00000000
2	.48039190-03	.24474571-03	.10475898-03	.51851717-05	.93569521-01	.49455843-01
3	.48062528-03	.24490670-03	.10480719-03	.10374396-04	.94234384-01	.98664136-01
4	.48101453-03	.24517527-03	.10488758-03	.15571732-04	.95341385-01	.14738464+00
5	.48156022-03	.24555178-03	.10500028-03	.20781259-04	.96888931-01	.19539113+00
6	.48226298-03	.24603673-03	.10514542-03	.26007074-04	.98874767-01	.24247721+00
7	.48312366-03	.24663077-03	.10532316-03	.31253307-04	.10129598+00	.28846080+00
8	.48414344-03	.24733472-03	.10553375-03	.36524157-04	.10414911+00	.33318695+00
9	.48532356-03	.24814953-03	.10577746-03	.41823766-04	.10743000+00	.37652935+00
10	.48666550-03	.24907630-03	.10605455-03	.47156497-04	.11113392+00	.41839078+00
11	.48817101-03	.25011633-03	.10636541-03	.52526663-04	.11525556+00	.45870120+00
12	.48984205-03	.25127100-03	.10671042-03	.57938682-04	.11978902+00	.49741631+00
13	.49168070-03	.25254195-03	.10709001-03	.63397053-04	.12472785+00	.53451476+00
14	.49368939-03	.25393093-03	.10750468-03	.68906375-04	.13006503+00	.56999511+00
15	.49567071-03	.25543986-03	.10795494-03	.74471342-04	.13579307+00	.60387272+00

LIST OF REFERENCES

1. J. Makhoul, "Linear Prediction: A Tutorial Review, "Proceedings of the IEEE, vol. 63, no. 4, April 1975, pp. 561-580.

2. T. J. Ulrych and T. N. Bishop, "Maximum Entropy Spectral Analysis and Autoregressive Decomposition," Reviews of Geophysics and Space Physics, vol. 13, no. 1, February 1975, pp. 183-200.

3. A. H. Nuttall, Spectral Analysis of a Univariate Process with Bad Data Points, via Maximum Entropy and Linear Predictive Techniques, NUSC Technical Report 5303, 26 March 1976.

4. J. P. Burg, "A New Analysis Technique for Time Series Data," NATO Advanced Study Institute on Signal Processing, Enschede, Netherlands, vol. 1, August 1968.

5. A. H. Nuttall, FORTRAN Program for Multivariate Linear Predictive Spectral Analysis, Employing Forward and Backward Averaging, NUSC Technical Document 5419, 19 May 1976.

6. R. H. Jones, "Multivariate Maximum Entropy Spectral Analysis," Applied Time Series Analysis Symposium, Tulsa Oklahoma, 14-15 May 1976. A revised version will appear in Applied Time Series Analysis, editor D. Findley, Academic Press, 1977.

7. P. Whittle, "On the Fitting of Multivariate Autoregressions, and the Approximate Canonical Factorization of a Spectral Density Matrix," Biometrika, vol. 50, 1963, pp. 129-134.

8. H. Akaike, "Block Toeplitz Matrix Inversion," SIAM J. Appl. Math., vol. 24, no. 2, March 1973, pp. 234-241.

9. G. Kreisselmeier, "A Solution of the Bilinear Matrix Equation AY+YB = -Q," SIAM J. Appl. Math., vol. 23, no. 3, Nov. 1972, pp. 334-338.

10. A. Jameson, "Solution of the Equation AX+XB = C by Inversion of an MxM or NxN Matrix," SIAM J. Appl. Math., vol. 16, no. 5, Sept. 1968, pp. 1020-1023.

11. C. S. Lu, "Solution of the Matrix Equation AX+XB = C," Electronics Letters, vol. 7, no. 8, April 22, 1971, pp. 185-186.

12. S. Barnett, "Remarks on Solution of AX+XB = C," Electronics Letters, vol. 7, no. 14, July 15, 1971, p. 385.

13. I. S. Pace and S. Barnett, "Comparison of Numerical Methods for Solving Liapunov Matrix Equations," Int. J. Control, vol. 15, no. 5, 1972, pp. 907-915.

14. R. E. Hartwig, "Resultants and the Solution of AX-XB = -C," SIAM J. Appl. Math., vol. 23, no. 1, July 1972, pp. 104-117.

15. V. Kučera, "The Matrix Equation AX+XB = C," SIAM J. Appl. Math., vol. 26, no. 1, Jan. 1974, pp. 15-25.

16. R. E. Hartwig, "AX-XB = C, Resultants and Generalized Inverses," SIAM J. Appl. Math., vol. 28, no. 1, Jan. 1975, pp. 154-183.

17. S. Barnett, "Simplification of Certain Linear Matrix Equations," IEEE Trans on Automatic Control, February 1976, pp. 115-116.

18. F. R. Gantmacher, "The Theory of Matrices," vol. I, Chelsea Publishing Co., N.Y., 1959.

19. H. Akaike, "A New Look at the Statistical Model Identification," IEEE Trans. on Automatic Control, vol. AC-19, no. 6, Dec. 1974, pp. 716-723.

20. A. H. Nuttall, "Trigonometric Smoothing and Interpolation of Sampled Complex Functions, via the FFT," NUSC Technical Memorandum TC-94-71, 19 April 1971.

21. J. F. Ferrie, G. C. Carter, and C. W. Nawrocki, "Availability of Markel's FFT Algorithm," NUSC Technical Memorandum TC-1-73, 15 January 1973.